

3

STIC-Biotech/ChemLib

145029

mg

From: Li, Bao-Qun
Sent: Monday, February 14, 2005 8:43 AM
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 20 IN APPLICATION SN.
09,664,363. THANKS.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

RECEIVED
FEB 16 2005
STIC

STAFF USE ONLY

Searcher: Arnold
Searcher Phone: 2-2532
Date Searcher Picked up: 2/15/05
Date Completed: 2/15/05
Searcher Prep/Rev. Time: 2/15/05
Online Time: _____

Type of Search

NA Sequence: # 1
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (u.s.p.),



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 145029

TO: Bao-Qun Li
Location: rem/3a24/3c18
Art Unit: 1648
Monday, February 28, 2005

Case Serial Number: 09/664363

From: Deirdre Arnold
Location: Biotech-Chem Library
REM 1A64
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

Search Notes

Please feel free to contact me if you have any questions or would like to amend the search.

Thank you for using STIC services.

Regards,
Deirdre Arnold

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 02:17:34 ; Search time 8554 Seconds
(without alignments)
11572.832 Million cell updates/sec

Title: US-09-664-363-20
Perfect score: 2043
Sequence: 1 TGGGAGGCGCTTCACAGC.....TGTGGCGGCGATGAGATGGGC 2043

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapect 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_secs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	2043	100.0	2043	6	A28153 PT-NANBH mr
2	2043	100.0	2043	6	A32201 NANBH PT p
3	2043	100.0	2043	6	ARI44049 Sequence
4	2043	100.0	3750	6	A28157 PT-NANBH mr
5	2043	100.0	3750	6	ARI44051 Sequence
6	2035	99.6	7065	6	A76577 Sequence 9
7	2022.2	99.0	3750	6	A32203 NANBH PT p
8	1804.6	88.3	9410	14	HPCKXR2 D50485 Hepatitis C
9	1802.6	88.2	9615	14	AB049088 Hepatitis C
10	1801.4	88.2	9587	14	AF333324 Hepatitis C
11	1798.2	88.0	9379	14	AF165060 Hepatitis C
12	1796.6	87.9	7863	6	AR064526 Sequence
13	1796.6	87.9	7863	6	AR091919 Sequence
14	1796.6	87.9	7863	6	AR064524 Sequence
15	1796.6	87.9	7863	6	AR064524 Sequence
16	1796.6	87.9	7917	6	AR005085 Sequence
17	1796.6	87.9	7917	6	AR064524 Sequence
18	1796.6	87.9	7917	6	AR091917 Sequence
19	1796.6	87.9	7917	6	ARI45566 Sequence

20	1796.6	87.9	9030	6	AR005076 Sequence
21	1796.6	87.9	9030	6	AR064515 Sequence
22	1796.6	87.9	9030	6	AR091908 Sequence
23	1796.6	87.9	9030	6	ARI45557 Sequence
24	1796.6	87.9	9416	6	AR005070 Sequence
25	1796.6	87.9	9416	6	AR064509 Sequence
26	1796.6	87.9	9416	6	AR091902 Sequence
27	1796.6	87.9	9416	6	AR145551 Sequence
28	1796.6	87.9	9416	6	E08315 cDNA encodi
29	1796.6	87.9	9416	14	HPCKXR
30	1795	87.9	9374	14	AF207753 Hepatitis C
31	1795	87.9	9386	14	AF165055 Hepatitis C
32	1795	87.9	9416	6	E10035 cDNA of non
33	1795	87.9	9573	14	AB049090 Hepatitis C
34	1793.4	87.8	9379	14	AF165059 Hepatitis C
35	1793.4	87.8	9379	14	HCVPOLYP
36	1793.4	87.8	9386	14	AF165056 Hepatitis C
37	1788.6	87.5	9379	14	AF207767 Hepatitis C
38	1788.6	87.5	9410	14	HPCKXR2
39	1787	87.5	7989	6	AR406043 Sequence
40	1787	87.5	7989	6	AX036255 Sequence
41	1787	87.5	7989	12	SSE242652 Hepatitis C
42	1787	87.5	8001	6	AR406042 Sequence
43	1787	87.5	8001	6	AX036254 Sequence
44	1787	87.5	8001	12	SSE242654 Hepatitis C
45	1787	87.5	8637	6	AR406041 Sequence

ALIGNMENTS

RESULT 1
A28153
LOCUS PT-NANBH mRNA fragment from patent GB2239245.
DEFINITION A28153
ACCESSION A28153
VERSION A28153.1 GI:1248636
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 2043)
AUTHORS
TITLE
JOURNAL
FEATURES
Source
Post-transfusal non-A non-B hepatitis viral polypeptides
Patent: GB 2239245-A 20 26-JUN-1991;
location/Qualifiers
1..2043
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="156/192"
<1..>2043
/codon_start=1
/transl_table=1
/product="PT-NANBH polypeptide"
/protein_id="CAA01930.1"
/db_xref="gi:1248637"

CDS

ORIGIN
Query Match 100.0%; Score 2043; DB 6; Length 2043;
Best Local Similarity 100.0%; Pred. No. 0;
PSWDMWKCLRLKPTLNGPPLLYLGAVNEVTLHPITKFTMACSADLEVTST
/translation="WEGVFTGLTHVDHAFLSQTKQAGNPFYLYAQTVCARQAP
WLVGVVLAALAAACLTGGSVVIIGRIISGRPAIVDPREVLQFDEMECSHLPY
IEGMOQLAEQKALGLQATKQAEAAAVESKSMIIVPRTCSNMHGTFTINAY
AGLSLPGNPAIASIMAFASVTSPLTOSTLLNIGGWAALAPASAPAFAG
IGCAVSGICGKLVLDIAGYCGVGAIVAFKMGSEMGSTEDIVLLPAIISPGA
LVGVVCAIILRRHYGREGVAVQMMNRILIAFSKRNHVSPIHYIPESDAARVOTIS
DLTTQLKRLHQMINEGCTSPGSGSWLRDWDNICTVLADFTKWLQKLPRLGVP
FFSCORGKYGWRGDIMQITCSGGAQTTHGVKSGSMIIVPRTCSNMHGTFTINAY
TTGCTPSPAPRALRVAEEVEYTRVGDHYVTSMTDVKCPQVPADEFPT
EVDGRLHRYVAPACKPLREVTQVGNLYVGSOLPCCEBPDAVAVTSLTDPDSHI
TAETAKRIIAGSPPSLASSASQLSAPSSKATYITQNDPFDADLIENLIMRHEMG"

Matches 2043: Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	TGGAGGGGCTTTCACAGGCTTCAACCCAGTGGATGCCACTTCTGTCCTCAAAACAAG	60	
DB	1	TGGAGGGGCTTTCACAGGCTTCAACCCAGTGGATGCCACTTCTGTCCTCAAAACAAG	60	
QY	61	CAGGAGAGAGAACTTCCCTTACCTGGTGGGCTACAGGCTTACGTGTGGCTTGGGGC	120	
DB	61	CAGGAGAGAGAACTTCCCTTACCTGGTGGGCTACAGGCTTACGTGTGGCTTGGGGC	120	
QY	121	CAGGCCCACTTCATCATGGATCAAAATGTGAAAGTGTCTATACGGCTAAAGCTACT	180	
DB	121	CAGGCCCACTTCATCATGGATCAAAATGTGAAAGTGTCTATACGGCTAAAGCTACT	180	
QY	181	CTGGCGGGGCAACACCTTGTGTATAGGCTTGGAGCCGTCACAAACAGAGTCAACCTC	240	
DB	181	CTGGCGGGGCAACACCTTGTGTATAGGCTTGGAGCCGTCACAAACAGAGTCAACCTC	240	
QY	241	ACACACCCCATPAAACCAATTCATCATGGATCATGTGATGCCAGCTGGAAGTTCGTACG	300	
DB	241	ACACACCCCATPAAACCAATTCATCATGGATCATGTGATGCCAGCTGGAAGTTCGTACG	300	
QY	301	AGCACCCTGGTCTGGTGGGCGGGGCTTTCAGCTCTGGCTGCGTATGCTTGACAACA	360	
DB	301	AGCACCCTGGTCTGGTGGGCGGGGCTTTCAGCTCTGGCTGCGTATGCTTGACAACA	360	
QY	361	GGCAGCGTGTCTATGTGGTGGATCATCTTGTCCGGGGCGGCTATGTTCCCGAC	420	
DB	361	GGCAGCGTGTCTATGTGGTGGATCATCTTGTCCGGGGCGGCTATGTTCCCGAC	420	
QY	421	AGGGAAGTCTTACAGAGAGTTTCATGATGAGAGAGTGGCGCTCGCACTCCCTTAC	480	
DB	421	AGGGAAGTCTTACAGAGAGTTTCATGATGAGAGAGTGGCGCTCGCACTCCCTTAC	480	
QY	481	ATCGAGCAGGAGATCAGCTCCGCGAGCAGTTCAGCAAAACCGCTGGGTTGCTGACG	540	
DB	481	ATCGAGCAGGAGATCAGCTCCGCGAGCAGTTCAGCAAAACCGCTGGGTTGCTGACG	540	
QY	541	ACAGCCACCAAGCAAGCGAGGCGGCTGCTCCGTTGGTGGAGTCCAAGTGGCGAGCCCTT	600	
DB	541	ACAGCCACCAAGCAAGCGAGGCGGCTGCTCCGTTGGTGGAGTCCAAGTGGCGAGCCCTT	600	
QY	601	GAGACCTTCTGGGCGAAACACATGTGGAACTTTCATCAGCGGATACAGTACTTACAGGC	660	
DB	601	GAGACCTTCTGGGCGAAACACATGTGGAACTTTCATCAGCGGATACAGTACTTACAGGC	660	
QY	661	TGTGCACTCTGCTGGGAACTCCCGGATGTGCATCATGTAGGCTTCAACAGCTCTGTC	720	
DB	661	TGTGCACTCTGCTGGGAACTCCCGGATGTGCATCATGTAGGCTTCAACAGCTCTGTC	720	
QY	721	ACTAGCCGCTCACCAACCAATCTACCTCTGCTTAAACATCTCGGGGGATGGGTAGCC	780	
DB	721	ACTAGCCGCTCACCAACCAATCTACCTCTGCTTAAACATCTCGGGGGATGGGTAGCC	780	
QY	781	GCCCAACTGCTCCCCCAGTGTCTTACGCTTTCGTAGGCGCGGCAATCTGCTGCTG	840	
DB	781	GCCCAACTGCTCCCCCAGTGTCTTACGCTTTCGTAGGCGCGGCAATCTGCTGCTG	840	
QY	841	GCTGTTGGAGCATAGGCTTGGGAAAGTGTGTGGACATCTTGGCGGGCTATGAGACA	900	
DB	841	GCTGTTGGAGCATAGGCTTGGGAAAGTGTGTGGACATCTTGGCGGGCTATGAGACA	900	
QY	901	GGAGTGGCAGGCGCTGTGGCTTTTAAAGTCAATGAGCGGAGAAATGCCCTCACCGAG	960	
DB	901	GGAGTGGCAGGCGCTGTGGCTTTTAAAGTCAATGAGCGGAGAAATGCCCTCACCGAG	960	
QY	961	GACCTGGTTAACTTACTCTCTGCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020	
DB	961	GACCTGGTTAACTTACTCTCTGCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020	
QY	1021	TGCGAGAGCATCTGCTGGGACGTTGGTCCAGGGGAGGGGGCTGTGAGTGAATGAAC	1080	
DB	1021	TGCGAGAGCATCTGCTGGGACGTTGGTCCAGGGGAGGGGGCTGTGAGTGAATGAAC	1080	

QY	1081	CGGCTGATAGCGTTGCGCTCGCGGGGTAAACCATGTTTCCCCACGCACTATGTGCCAGAG	1140	
DB	1081	CGGCTGATAGCGTTGCGCTCGCGGGGTAAACCATGTTTCCCCACGCACTATGTGCCAGAG	1140	
QY	1141	AGCGACGCGAGCAGTGTCACTGATTCCTCCAGCTTAACTACACCAACTGTTG	1200	
DB	1141	AGCGACGCGAGCAGTGTCACTGATTCCTCCAGCTTAACTACACCAACTGTTG	1200	
QY	1201	AAAGGCTTCAACAGTGAATTAACAGAGACTGCTCACGCCCTGCTCGGCTCGTACTA	1260	
DB	1201	AAAGGCTTCAACAGTGAATTAACAGAGACTGCTCACGCCCTGCTCGGCTCGTACTA	1260	
QY	1261	AGGATGTTTGGGACTGATATGCAACATTTTGGTGACTTCAAGACTTGCTCCAGTCC	1320	
DB	1261	AGGATGTTTGGGACTGATATGCAACATTTTGGTGACTTCAAGACTTGCTCCAGTCC	1320	
QY	1321	AAGCTCTGCGCGATPAAACCGGGAGTCCCTTTTCTCATGCGCAACGTGGGTAAAGGGG	1380	
DB	1321	AAGCTCTGCGCGATPAAACCGGGAGTCCCTTTTCTCATGCGCAACGTGGGTAAAGGGG	1380	
QY	1381	GTTGGCGGGGAGACGGCATATGACACACTGCTCATGTGAGCAACAGATCACCGGA	1440	
DB	1381	GTTGGCGGGGAGACGGCATATGACACACTGCTCATGTGAGCAACAGATCACCGGA	1440	
QY	1441	CATGTCAAAAACGTTTCATAGAGATCGTTGGGCTTAAGACTGTGATTAACATGTGCAT	1500	
DB	1441	CATGTCAAAAACGTTTCATAGAGATCGTTGGGCTTAAGACTGTGATTAACATGTGCAT	1500	
QY	1501	GGAACTTCCCATCAACGCAATPACACACGCGGCGCTGACAGCCCTCCCGAGCGCAAC	1560	
DB	1501	GGAACTTCCCATCAACGCAATPACACACGCGGCGCTGACAGCCCTCCCGAGCGCAAC	1560	
QY	1561	TATTCAGAGGCGCTGTGGCGGGTGGCTGAGAGATACGTGAGATTACCGGGTGGGG	1620	
DB	1561	TATTCAGAGGCGCTGTGGCGGGTGGCTGAGAGATACGTGAGATTACCGGGTGGGG	1620	
QY	1621	GATTTCCACTACGTGACGAGCATGACCACTGACAAACGTAAATGCGCCGTGCAAGTTCCA	1680	
DB	1621	GATTTCCACTACGTGACGAGCATGACCACTGACAAACGTAAATGCGCCGTGCAAGTTCCA	1680	
QY	1681	GCCCCGGAATCTTACAGAAAGTGAATGGGGTGGCGGCTGCAACAGTACGCTCCGCGTGC	1740	
DB	1681	GCCCCGGAATCTTACAGAAAGTGAATGGGGTGGCGGCTGCAACAGTACGCTCCGCGTGC	1740	
QY	1741	AAACCTTCTCTACGAGGAGAGGTCACTTCCAGATCGGGCTTCAACCAATACCTGTTGGG	1800	
DB	1741	AAACCTTCTCTACGAGGAGAGGTCACTTCCAGATCGGGCTTCAACCAATACCTGTTGGG	1800	
QY	1801	TGCGAGCTCCCATGCGAGCCGGAACCGGATGTAGAGTGTCTCATATGCTCACCGAC	1860	
DB	1801	TGCGAGCTCCCATGCGAGCCGGAACCGGATGTAGAGTGTCTCATATGCTCACCGAC	1860	
QY	1861	CCCTCCCATCAACAGCAGAGCGGCTTAAAGCGAGGCTGAGCGAGGGGTCTCCCCCTCC	1920	
DB	1861	CCCTCCCATCAACAGCAGAGCGGCTTAAAGCGAGGCTGAGCGAGGGGTCTCCCCCTCC	1920	
QY	1921	TTGGCAGCTTTCAGTACGAGTGTCTGCGCTTCTCGAAGCGACATATTAAC	1980	
DB	1921	TTGGCAGCTTTCAGTACGAGTGTCTGCGCTTCTCGAAGCGACATATTAAC	1980	
QY	1981	CAAAATGACTTCCAGACGCTGACCTCATGACAGGCAACCTCTGTGCGGACATGATG	2040	
DB	1981	CAAAATGACTTCCAGACGCTGACCTCATGACAGGCAACCTCTGTGCGGACATGATG	2040	
QY	2041	GGC 2043		
DB	2041	GGC 2043		

RESULT 2
A32201 A32201 LOCUS A32201 2043 bp DNA linear PAT 10-DEC-1996

QY 1621 GATTTCACCTACGTGACGAGCATGACCACTGACAAAGTAAATGCGCCGTGCGAGGTTCCA 1680
| | | | |
DB 1621 GATTTCACCTACGTGACGAGCATGACCACTGACAAAGTAAATGCGCCGTGCGAGGTTCCA 1680
| | | | |
QY 1681 GCCCCCGAATTTTTCACAGAAATGGATGGGGTGGCGGCTGACAGAGTACGCTCCGGCGTGC 1740
| | | | |
DB 1681 GCCCCCGAATTTTTCACAGAAATGGATGGGGTGGCGGCTGACAGAGTACGCTCCGGCGTGC 1740
| | | | |
QY 1741 AAACCTCTCCATCGGAGAGAGGTGACATTTCCAGATCGGGCTCAACCAATATCTGGTGGG 1800
| | | | |
DB 1741 AAACCTCTCTACGGAGAGAGGTGACATTTCCAGATCGGGCTCAACCAATATCTGGTGGG 1800
| | | | |
QY 1801 TCGAGCTTCCATGCGAGGCCGCAACCGGATGTAGACAGTCTCATTCCATGCTCAACGAC 1860
| | | | |
DB 1801 TCGAGCTTCCATGCGAGGCCGCAACCGGATGTAGACAGTCTCATTCCATGCTCAACGAC 1860
| | | | |
QY 1861 CCTTCCCATCATCAGCAGAGACGGCTAAGCGCAGGCTGGCCAGGGGGTCTCCCTCC 1920
| | | | |
DB 1861 CCTTCCCATCATCAGCAGAGACGGCTAAGCGCAGGCTGGCCAGGGGGTCTCCCTCC 1920
| | | | |
QY 1921 TTGGCAGCTTCTTACGCTGACGATTTGTCTGCGCCTTCTCTCGAAGGCGCATATTAACC 1980
| | | | |
DB 1921 TTGGCAGCTTCTTACGCTGACGATTTGTCTGCGCCTTCTCTCGAAGGCGCATATTAACC 1980
| | | | |
QY 1981 CAAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTTCCTGTGGCGGATGAGATG 2040
| | | | |
DB 1981 CAAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTTCCTGTGTGGCGGATGAGATG 2040
| | | | |
QY 2041 GGC 2043
| | |
DB 2041 GGC 2043
| | |

RESULT 3
AR144049 2043 bp DNA linear PAT 08-AUG-2001
LOCUS AR144049
DEFINITION Sequence 20 from patent US 6210675.
ACCESSION AR144049
VERSION AR144049.1 GI:15105916
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2043)
AUTHORS Highfield, P.Edmund., Rodgers, B.,Colin., Tedder, R.,Seton. and
TITLE PT-NANB hepatitis B polypeptides
JOURNAL Patent: US 6210675-A 20 03-APR-2001;
FEATURES
source
1..2043
/organism="unknown"
/mol_type="unasigned DNA"

Query Match 100.0%; Score 2043; DB 6; Length 2043;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGAGGGGCGTCTTCAAGGCTTCAACCAAGTGGATGCCACTTCTGTCCAAACAAG 60
| | | | |
DB 1 TGGGAGGGGCGTCTTCAAGGCTTCAACCAAGTGGATGCCACTTCTGTCCAAACAAG 60
| | | | |
QY 61 CAGGAGAGAGAACTTCCCTTACCTGGTGGGCTACAGGCTACCTGTGGGCTTGGGGC 120
| | | | |
DB 61 CAGGAGAGAGAACTTCCCTTACCTGGTGGGCTACAGGCTACCTGTGGGCTTGGGGC 120
| | | | |
QY 121 CAGGCCCCACCTTCATCATGGATCAAAATGTGGAATGTCTCATAGCGCTAAAGCTTACT 180
| | | | |
DB 121 CAGGCCCCACCTTCATCATGGATCAAAATGTGGAATGTCTCATAGCGCTAAAGCTTACT 180
| | | | |
QY 181 CTGCGCGGGCCAAACACCTTGTGTATAGGCTGGGAGCGGTCCAAAAGAGGTCAACCTTC 240
| | | | |

DB 181 CTGCGCGGGCCAAACACCTTGTGTATAGGCTGGGAGCGGTCCAAAAGAGGTCAACCTTC 240
| | | | |
QY 241 ACACACCCCATTAACCAATTTTCATATGCGATGATGTGACCGCATCTGGAGGTGTCACG 300
| | | | |
DB 241 ACACACCCCATTAACCAATTTTCATATGCGATGATGTGACCGCATCTGGAGGTGTCACG 300
| | | | |
QY 301 AGCACCTGGGTGCTGGTGGGGGGTCTTTCAGCTCTGGCTGGGTATTTGCTTGAACA 360
| | | | |
DB 301 AGCACCTGGGTGCTGGTGGGGGGTCTTTCAGCTCTGGCTGGGTATTTGCTTGAACA 360
| | | | |
QY 361 GGCAGCGGTGATCATTTGTGGATAGATCATCTTGTCCGGGGCGGCTATTGTTCCGAC 420
| | | | |
DB 361 GGCAGCGGTGATCATTTGTGGATAGATCATCTTGTCCGGGGCGGCTATTGTTCCGAC 420
| | | | |
QY 421 AGGGAAGTCTCTACAGAGATTGATAGATGGAAGAGTGCCTGACCTCCCTTAC 480
| | | | |
DB 421 AGGGAAGTCTCTACAGAGATTGATAGATGGAAGAGTGCCTGACCTCCCTTAC 480
| | | | |
QY 481 ATCGAGCGGGAATGCACTGCTGCGAGAGATTCAAGCAAAAAGCGCTGGGTTGCTGGAG 540
| | | | |
DB 481 ATCGAGCGGGAATGCACTGCTGCGAGAGATTCAAGCAAAAAGCGCTGGGTTGCTGGAG 540
| | | | |
QY 541 ACAGCCACCAAGCAAGCGGAGCGCTGCTCCCGTGTGAGTCCAAATGGCAGCCCTT 600
| | | | |
DB 541 ACAGCCACCAAGCAAGCGGAGCGCTGCTCCCGTGTGAGTCCAAATGGCAGCCCTT 600
| | | | |
QY 601 GAGACCTTCTGGGGGAAAACATGTGAACTTCAATCAGCGGGAATACATTAAGCAGGC 660
| | | | |
DB 601 GAGACCTTCTGGGGGAAAACATGTGAACTTCAATCAGCGGGAATACATTAAGCAGGC 660
| | | | |
QY 661 TTGTCACTCTGCTGGGGAATCCGCGATTTGATCATCTGATGAGGCTTCAAGCTCTGTC 720
| | | | |
DB 661 TTGTCACTCTGCTGGGGAATCCGCGATTTGATCATCTGATGAGGCTTCAAGCTCTGTC 720
| | | | |
QY 721 ACTAGCCGCTCACACCAATTAACCTCCGCTTAATCATCTGAGGGGATGAGGTACCC 780
| | | | |
DB 721 ACTAGCCGCTCACACCAATTAACCTCCGCTTAATCATCTGAGGGGATGAGGTACCC 780
| | | | |
QY 781 GCCCAACTCGTCCCGCCAGTGTCTGCTTCACTTTCGTAGGCGCGGCAATTTGCTGTCG 840
| | | | |
DB 781 GCCCAACTCGTCCCGCCAGTGTCTGCTTCACTTTCGTAGGCGCGGCAATTTGCTGTCG 840
| | | | |
QY 841 GCTGTTGCAAGATGAGCTTGGGAGGTGCTTGTGACATCTTGGCGGGCTATGAGACA 900
| | | | |
DB 841 GCTGTTGCAAGATGAGCTTGGGAGGTGCTTGTGACATCTTGGCGGGCTATGAGACA 900
| | | | |
QY 901 GGAATGGAAGCGGCTGTGGCTTTAAGGATGAGAGCGGGAATGCCCTCCACGAG 960
| | | | |
DB 901 GGAATGGAAGCGGCTGTGGCTTTAAGGATGAGAGCGGGAATGCCCTCCACGAG 960
| | | | |
QY 961 GACTGTGTTAACTTACTCTGCACTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
| | | | |
DB 961 GACTGTGTTAACTTACTCTGCACTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
| | | | |
QY 1021 TGGCAGAGATATCTGCTGCGACGCTGGTCCAGGGGAGGGGCTGTGCAATGATGAAC 1080
| | | | |
DB 1021 TGGCAGAGATATCTGCTGCGACGCTGGTCCAGGGGAGGGGCTGTGCAATGATGAAC 1080
| | | | |
QY 1081 CGGCTGATAGAGTTCGCTCGGGGGTAAACAATGTTTCCCAAGCAGCATATGTGCAGAG 1140
| | | | |
DB 1081 CGGCTGATAGAGTTCGCTCGGGGGTAAACAATGTTTCCCAAGCAGCATATGTGCAGAG 1140
| | | | |
QY 1141 AGCGAGCGCGAGCAGTGTCTCAGATCTCTCCAGCTTACTATGACCAACCTGTTG 1200
| | | | |
DB 1141 AGCGAGCGCGAGCAGTGTCTCAGATCTCTCCAGCTTACTATGACCAACCTGTTG 1200
| | | | |
QY 1201 AAGAGCTTCAACCAATGATTAACAGAGATGCTTCAAGCCCTGCTCGGCTGCTGCTG 1260
| | | | |
DB 1201 AAGAGCTTCAACCAATGATTAACAGAGATGCTTCAAGCCCTGCTCGGCTGCTGCTG 1260
| | | | |
QY 1261 AGGATGTTTGGGATCTGATATGACAGTTTGGTGTACTTCAAACTGGGTCCAGTCC 1320
| | | | |
DB 1261 AGGATGTTTGGGATCTGATATGACAGTTTGGTGTACTTCAAACTGGGTCCAGTCC 1320
| | | | |

QY 1321 AAGCTCTGCGCGGATTAACGGAGTCCCTTTTCTCATGCAAGTGGGTACAAGGG 1380
 DB 1321 AAGCTCTGCGCGGATTAACGGAGTCCCTTTTCTCATGCAAGTGGGTACAAGGG 1380
 QY 1381 GTCTGGCGGGAGACGGCATCATGACACCACTGCTCATGTGGACACAGATCACCGA 1440
 DB 1381 GTCTGGCGGGAGACGGCATCATGACACCACTGCTCATGTGGACACAGATCACCGA 1440
 QY 1441 CATGTCAAAAAGGTTTCCATGAGAGTCTTTGGGCTTAAGACCTGTGTAACTATGTGGCAT 1500
 DB 1441 CATGTCAAAAAGGTTTCCATGAGAGTCTTTGGGCTTAAGACCTGTGTAACTATGTGGCAT 1500
 QY 1501 GGAACATTTCCCATCAACGATACACCAAGGGGCTTGCACGGCTTCCCAAGGCGCAAC 1560
 DB 1501 GGAACATTTCCCATCAACGATACACCAAGGGGCTTGCACGGCTTCCCAAGGCGCAAC 1560
 QY 1561 TATTCAGAGGCGCTGTGGCGGCTGCTGAGAGTACGTGAGGTTACGCGGATGGG 1620
 DB 1561 TATTCAGAGGCGCTGTGGCGGCTGCTGAGAGTACGTGAGGTTACGCGGATGGG 1620
 QY 1621 GATTTCCACTACGTGACGAGCATGACCACTGACAACTAAATGCCGTGCCAGTTCCA 1680
 DB 1621 GATTTCCACTACGTGACGAGCATGACCACTGACAACTAAATGCCGTGCCAGTTCCA 1680
 QY 1681 GCGCCCGAATTTCTTCAAGAGTGAATGGGGTGGCGTGCACAGTACGCTCCGCGGTGC 1740
 DB 1681 GCGCCCGAATTTCTTCAAGAGTGAATGGGGTGGCGTGCACAGTACGCTCCGCGGTGC 1740
 QY 1741 AAACCTCTCTCAAGGAGAGGTACATTCAGAGTGGGCTCAACCAATACCTGTGGG 1800
 DB 1741 AAACCTCTCTCAAGGAGAGGTACATTCAGAGTGGGCTCAACCAATACCTGTGGG 1800
 QY 1801 TGGCAGCTCTCCATGCGAGCCCAACCGGATGTAGAGTGTCTCATTCATGCTCACCGAC 1860
 DB 1801 TGGCAGCTCTCCATGCGAGCCCAACCGGATGTAGAGTGTCTCATTCATGCTCACCGAC 1860
 QY 1861 CCCTCCCATCATCAAGCAGAGAGCGGTAAGCGGAGCTGGCCAGGGGGTCTCCCGCTCC 1920
 DB 1861 CCCTCCCATCATCAAGCAGAGAGCGGTAAGCGGAGCTGGCCAGGGGGTCTCCCGCTCC 1920
 QY 1921 TTGGCAGCTCTTCAAGTACGCACTGATGTGTGCGCTTCTCGAAGGCGACATACATTAAC 1980
 DB 1921 TTGGCAGCTCTTCAAGTACGCACTGATGTGTGCGCTTCTCGAAGGCGACATACATTAAC 1980
 QY 1981 CAAATGATCTTCCAGACGCTGACCTCATGAGGCGCACTCTGTGGCGGATGAGATG 2040
 DB 1981 CAAATGATCTTCCAGACGCTGACCTCATGAGGCGCACTCTGTGGCGGATGAGATG 2040
 QY 2041 GGC 2043
 DB 2041 GGC 2043

RESULT 4
 LOCUS A28157 3750 bp DNA linear PART 04-AUG-1995
 DEFINITION FT-NANBH mRNA fragment from patent GB239245.
 ACCESSION A28157 GI:1248640
 VERSION A28157.1 GI:1248640
 KEYWORDS
 SOURCE
 ORGANISM
 OTHER SEQUENCES; artificial sequences.
 REFERENCE
 TITLE Post-transferrin non-A non-B hepatitis viral polypeptides
 JOURNAL Patent: GB 239245-A 22-26-JUN-1991;
 FEATURES
 1.3750 Location/Qualifiers
 source

/organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"

CDS
 /note="contig formed by cdna clones from 3' end"
 <1..>3750
 /codon_start=1
 /transl_table=11
 /product="FT-NANBH polypeptide"
 /protein_id="CAA01932.1"
 /db_xref="GI:1248641"
 /translation="MEGVFTGLTHVDHFLSQTQKAGDNFYLVAVQATVCARAQAP
 PSMQMMKTLRLKPTLNGPPLRLYLRAICAVQNEVTLTHPILFKIEMACNADLEVST
 WLVGVLAALAAVYLTGTSVIVTRILLSGRPALVPRREVLYQEDMEBCASHLPY
 ISQNGVLAEOFKQKGLQITKQAEAPAVESKMRLETFMAKMNPTLSGTOYL
 AGLSLPGNPLASAFASVSPILTQSTLLNLI LGQWVAQOLA PSAAPAFVAG
 IAGAVGSIQGLKVLVDILAGVAGVAGLVFLFKYMGSGMPESTEDLVNLPILPGA
 LVGVGCAILIERHVGPGGAVQOMNRIAPFASRNHNSPTHYVESDAAAVTOILS
 DLITQILKRLHWNEDCSPTSGSMRLCTVLADKFTMLQSLRLRGLVP
 PFSQGRYGVVRGGINQTTSCCAQITGHVKNSSMIVGKTSNMHGFPPINAY
 TTGCTPSPAPVSRALNRVAEEVEVTRVDFHYVISMITDNKPCQVAPPEFT
 EVDGVRHLRYAPACKPLREBEVTFQVGNQYLVGSQLPCPEPDDVAVLISMLTDESHI
 TAEAKRLARGSPSLASASASAKVYITQNDFPDADILEANLIMRHMG
 DITRVESENKIVILSPDLRAEBDEREVSPAEILIRKSKRPPAMPAPADVNPPL
 LESKAPDVPPVHVCPLPPTKPTPIPEPRKRTVVLTESTVSSALAEIARFGSS
 EBSAVDSGTATAPPPQSDDGAGSDVESYSMPLEBEGPPLSDSMSIVSEBAG
 EDVCCSMSYMTGALITPCAESKLPINALSNLILRHNNVATITSRASQOKKV
 TFDRLQILDDHYQDVLKEMKAVSTVKAQLSVEBCKLTPHSAKSKRFGYAKOVN
 LSSKAIINHRSVWEDLEBDETETPIDITIMAKNEVFCVOPERGRRPARLIVPDIGVR
 VCEKALVDVSTLPOAVMGSSYGFQYSGRVEFLVNAWMSKPTMGPAYDTRFDS
 TYTENDIYVESITVCCDLAPARQALISITGRVLYIGELTMSKONGCYRRCRASGV
 LTTSGNTLITCTIKASACRAKLDCTMLVCGDELVIYCSAGTQEDBASLRVTEA
 MTRYAPDPDPQPEYDLELITSSNSVVAHDAAGKRVYLYLRDP"

ORIGIN

Query Match 100.0%; Score 2043; DB 6; Length 3750;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGAGGCGTCTTCAAGAGCTCAACCACTGATGCCACTTCTGTCACCAACAAG 60
 DB 1 TGGAGAGGCGTCTTCAAGAGCTCAACCACTGATGCCACTTCTGTCACCAACAAG 60
 QY 61 CAGGACAGAGACAATTCCTCCCTTACTGCTGGGCTACAGGCTACTGTGCGCTAGGGC 120
 DB 61 CAGGACAGAGACAATTCCTCCCTTACTGCTGGGCTACAGGCTACTGTGCGCTAGGGC 120
 QY 121 CAGGCCCACTCCATCATGAGATCAAAATGTGAAAGTGTCTCATGCGGTTAAAGCTACT 180
 DB 121 CAGGCCCACTCCATCATGAGATCAAAATGTGAAAGTGTCTCATGCGGTTAAAGCTACT 180
 QY 181 CTGCGGCGGCAACACCTTGTCTGTATAGCTGGAGCCGTCACAAACGAGGTCAACCTTC 240
 DB 181 CTGCGGCGGCAACACCTTGTCTGTATAGCTGGAGCCGTCACAAACGAGGTCAACCTTC 240
 QY 241 ACACACCCCATTAACCAATTCATCATGAGCATGTCAAGCCGACTGAGAGTCTGACAG 300
 DB 241 ACACACCCCATTAACCAATTCATCATGAGCATGTCAAGCCGACTGAGAGTCTGACAG 300
 QY 301 AGCAGCTGGTGTCTGGTGGCGGGGCTTGGAGCTGGCGGTGCGATGTGCTTGAACAAC 360
 DB 301 AGCAGCTGGTGTCTGGTGGCGGGGCTTGGAGCTGGCGGTGCGATGTGCTTGAACAAC 360
 QY 361 GGCAGGTGTGATCTGTGGGTAGATCATCTTGTCCGGCGGCGGCTATTTGTCCGAC 420
 DB 361 GGCAGGTGTGATCTGTGGGTAGATCATCTTGTCCGGCGGCGGCTATTTGTCCGAC 420
 QY 421 AAGGAAGTCTCTTACCAAGAGTTGATGATGAGAAAGTGGCGGTGCGACTCCCTTAC 480
 DB 421 AAGGAAGTCTCTTACCAAGAGTTGATGATGAGAAAGTGGCGGTGCGACTCCCTTAC 480
 QY 481 ATGAGAGGGAATGAGCTGCGAGCATCAAGCAAAAAGCGTGGGTGCTGACG 540
 DB 481 ATGAGAGGGAATGAGCTGCGAGCATCAAGCAAAAAGCGTGGGTGCTGACG 540
 QY 541 ACAGCCACCAACCAAGCGAGGCGCTGCTCCGCTGTGAGACTCCAAGTGGCAGCCCTT 600
 DB 541 ACAGCCACCAACCAAGCGAGGCGCTGCTCCGCTGTGAGACTCCAAGTGGCAGCCCTT 600

Db 541 ACAGCCACCAAGCAAGCGAGGCGCGCTGCTCCCGTGTGAGTCCAAAGTGGCAGCCCTT 600
 Qy 601 GAGACCTTCTGGGGGAAACACATGTGAATTCATCAGCGGATACAGTACTTACAGGC 660
 Db 601 GAGACCTTCTGGGGGAAACACATGTGAATTCATCAGCGGATACAGTACTTACAGGC 660
 Qy 661 TTGTCACTCTGCTGGGAAATCCCGGATGTCATCATGATGGCTTACAGCCTCTGTG 720
 Db 661 TTGTCACTCTGCTGGGAAATCCCGGATGTCATCATGATGGCTTACAGCCTCTGTG 720
 Qy 721 ACTACCCGCTACCAACCAATCTACCCCTGCTTAACTCTGCGGGGATGGGTAGGC 780
 Db 721 ACTACCCGCTACCAACCAATCTACCCCTGCTTAACTCTGCGGGGATGGGTAGGC 780
 Qy 781 GCCCAACTGCTCCCCCAGTGTCTTACGCTTTCATGAGCGCGGCAATGCTGTG 840
 Db 781 GCCCAACTGCTCCCCCAGTGTCTTACGCTTTCATGAGCGCGGCAATGCTGTG 840
 Qy 841 GCTGTGGCAGATAGGCTTGGGAAAGTGTGTGAGATCTTGGCGGGCTATGAGCA 900
 Db 841 GCTGTGGCAGATAGGCTTGGGAAAGTGTGTGAGATCTTGGCGGGCTATGAGCA 900
 Qy 901 GGAAGTGGCAGCGCGCTGTGGCTTTAAGTCATGAGCGGGAATGCCCTCCACCGAG 960
 Db 901 GGAAGTGGCAGCGCGCTGTGGCTTTAAGTCATGAGCGGGAATGCCCTCCACCGAG 960
 Qy 961 GACCTGTAACTTACCTCCCTGCAATCTCTCTCTGTGTGCTGTGTGCTGTGCTGTG 1020
 Db 961 GACCTGTAACTTACCTCCCTGCAATCTCTCTCTGTGTGCTGTGTGCTGTGCTGTG 1020
 Qy 1021 TGGCAGCGATATCTGCTGTGCACTGTGGTCCAGGGGAGGGGCTGTGATGATGAAC 1080
 Db 1021 TGGCAGCGATATCTGCTGTGCACTGTGGTCCAGGGGAGGGGCTGTGATGATGAAC 1080
 Qy 1081 CGGCGATAGAGCTTGGCTGTGGGGGTAACATGTTTCCCAACGACATATGAGCAAG 1140
 Db 1081 CGGCGATAGAGCTTGGCTGTGGGGGTAACATGTTTCCCAACGACATATGAGCAAG 1140
 Qy 1141 AGCGAGCGCGAGCACTGTCTCATCTGATCTCTCGACCTTACTATCAACCACTGTG 1200
 Db 1141 AGCGAGCGCGAGCACTGTCTCATCTGATCTCTCGACCTTACTATCAACCACTGTG 1200
 Qy 1201 AAGAGCTTCCACCAAGTGAATTAACAAGACTGTCTCCAGCCTGTCTGGCTGTGCTA 1260
 Db 1201 AAGAGCTTCCACCAAGTGAATTAACAAGACTGTCTCCAGCCTGTCTGGCTGTGCTA 1260
 Qy 1261 AAGGAGTGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 Db 1261 AAGGAGTGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 Qy 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGTACAAGGG 1380
 Db 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGTACAAGGG 1380
 Qy 1381 GTCTGGCGGGAGACGGCATCATGACACACCTGTCTCATGAGCAAGATCACCGGA 1440
 Db 1381 GTCTGGCGGGAGACGGCATCATGACACACCTGTCTCATGAGCAAGATCACCGGA 1440
 Qy 1441 CATGTCAAAAAGCTTCCATGAGATGATGAGATGATGAGATGATGAGATGATGAGAT 1500
 Db 1441 CATGTCAAAAAGCTTCCATGAGATGATGAGATGATGAGATGATGAGATGATGAGAT 1500
 Qy 1501 GGAACATTTCCCATCAACGATTAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
 Db 1501 GGAACATTTCCCATCAACGATTAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
 Qy 1561 TATTCAGGGGCGCTGTGTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
 Db 1561 TATTCAGGGGCGCTGTGTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
 Qy 1621 GATTTCACTAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
 Db 1621 GATTTCACTAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680

Qy 1681 GCCCCGAAATTTCTTCAAGAAAGTGAATGGGTGGGCTGTGACAGGTACGCTTCCGCGTGC 1740
 Db 1681 GCCCCGAAATTTCTTCAAGAAAGTGAATGGGTGGGCTGTGACAGGTACGCTTCCGCGTGC 1740
 Qy 1741 AAACCTCTTCAAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 Db 1741 AAACCTCTTCAAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 Qy 1801 TCGAGCTCCCATGCGAGCCCGAAACCGGATGATGATGATGATGATGATGATGATGATGAT 1860
 Db 1801 TCGAGCTCCCATGCGAGCCCGAAACCGGATGATGATGATGATGATGATGATGATGATGAT 1860
 Qy 1861 CCTTCCCATGACAGAGACGCTTAAGGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
 Db 1861 CCTTCCCATGACAGAGACGCTTAAGGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
 Qy 1921 TTGGCAGCTCTTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
 Db 1921 TTGGCAGCTCTTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
 Qy 1981 CAAATGACTTCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Db 1981 CAAATGACTTCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Qy 2041 GGC 2043
 Db 2041 GGC 2043
 RESULT 5
 ARI44051
 LOCUS ARI44051 3750 bp DNA linear PAT 08-AUG-2001
 DEFINITION Sequence 22 from patent US 6210675.
 ACCESSION ARI44051
 VERSION ARI44051.1 GI:15105918
 KEYWORDS
 SOURCE
 ORGANISM
 UNKNOWN.
 unclassified.
 1 (bases 1 to 3750)
 REFERENCE
 Highfield, P. Edmund., Rodgers, B. Colin., Tedder, R. Seton. and
 AUTHORS Barbara, J. Anthony, James.
 TITLE PT-NAM hepatitis polypeptides
 JOURNAL Patent: US 6210675-A 22 03-APR-2001;
 FEATURES
 Location/Qualifiers
 source 1..3750
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Query Match 100.0%; Score 2043; DB 6; Length 3750;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TGGAGGGCGTCTTCAAGGCTTCAACCAAGTGAATGTCCTTCTGTTCCAAACAAG 60
 Db 1 TGGAGGGCGTCTTCAAGGCTTCAACCAAGTGAATGTCCTTCTGTTCCAAACAAG 60
 Qy 61 CAGGAGAGAACAAATTTCCCTAAGTGTGGGTGATGAGGCTGATGAGGCTGAGGCTGAGGCT 120
 Db 61 CAGGAGAGAACAAATTTCCCTAAGTGTGGGTGATGAGGCTGATGAGGCTGAGGCTGAGGCT 120
 Qy 121 CAGGCCCATCTCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 Db 121 CAGGCCCATCTCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 Qy 181 CTGCGCGGCGCAACACCTTGTGTATAGGCTGTGAGCGCTTCAACAAGAGTCAACCTTC 240
 Db 181 CTGCGCGGCGCAACACCTTGTGTATAGGCTGTGAGCGCTTCAACAAGAGTCAACCTTC 240
 Qy 241 ACACACCCATTAACCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 Db 241 ACACACCCATTAACCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300

Db	2414	ACACAGCCCCATAACCAAAATTAATCATATGGCATCATGACGCCGACTCGGAAGGTGCTCACG	300
QY	301	AGCACTCTGGGTGCTGTGTGGGCGGGGCTCTTGACAGCTCTGGCTGCTATTTGCTTGACAACA	360
Db	301	AGCACTCTGGGTGCTGTGTGGGCGGGGCTCTTGACAGCTCTGGCTGCTATTTGCTTGACAACA	360
QY	361	GGCAGCGTGGTCAATTGTGGGTAGATATATCTGTGTCGGGGCGGCGGCTATTTGCTCGGAC	420
Db	361	GGCAGCGTGGTCAATTGTGGGTAGATATATCTGTGTCGGGGCGGCGGCTATTTGCTCGGAC	420
QY	421	AGGGAAGTCTCTACACAGAGTTCATGATGATGAAGAAGTGGCGCTCCACACTCCCTTAC	480
Db	421	AGGGAAGTCTCTACACAGAGTTCATGATGATGAAGAAGTGGCGCTCCACACTCCCTTAC	480
QY	481	ATCGAGCAGGGAAATCACACTCGCGGAGCAATTCAAGCAAAAAGCGCTCGGGTGTCTGACG	540
Db	481	ATCGAGCAGGGAAATCACACTCGCGGAGCAATTCAAGCAAAAAGCGCTCGGGTGTCTGACG	540
QY	541	ACAGGCACCAAGCAAGCGGAGGGCGGCTCTCCCGGTGTGGAGTCCAAATGGCGAGGCCCTT	600
Db	541	ACAGGCACCAAGCAAGCGGAGGGCGGCTCTCCCGGTGTGGAGTCCAAATGGCGAGGCCCTT	600
QY	601	GAGACTTCTGGGCGGAAAACAATGTGAACTTATCACGCGGATACAGTACTTATAGCAGGC	660
Db	601	GAGACTTCTGGGCGGAAAACAATGTGAACTTATCACGCGGATACAGTACTTATAGCAGGC	660
QY	661	TTGTCTCACTGTGCTGTGGGAATCCGCGAATTCATCATGTATGTGCTTACAGCCTCTGTC	720
Db	661	TTGTCTCACTGTGCTGTGGGAATCCGCGAATTCATCATGTATGTGCTTACAGCCTCTGTC	720
QY	721	ACTAGCCCGGCTTACCAACCAATCTTACCTCTGCTTAACTATCTGTGGGGGATGGGTAGCC	780
Db	721	ACTAGCCCGGCTTACCAACCAATCTTACCTCTGCTTAACTATCTGTGGGGGATGGGTAGCC	780
QY	781	GCCCAACTCGCTCCCCCGGAGTGTGCTTCAAGCTTTCGTAAGGCGCGGCAATGTGTGTGCG	840
Db	781	GCCCAACTCGCTCCCCCGGAGTGTGCTTCAAGCTTTCGTAAGGCGCGGCAATGTGTGTGCG	840
QY	841	GCTGTGGCAGCATTAAGGCTTGTGGAAAGTGTCTTGAGCATCTTGTGGCGGGCTATGAGACA	900
Db	841	GCTGTGGCAGCATTAAGGCTTGTGGAAAGTGTCTTGAGCATCTTGTGGCGGGCTATGAGACA	900
QY	901	GGAATGGCAGGCGCGCTGTGGCCTTTAAAGTCATGAGCGCGGAATGCCCTCCACCGAG	960
Db	901	GGAATGGCAGGCGCGCTGTGGCCTTTAAAGTCATGAGCGCGGAATGCCCTCCACCGAG	960
QY	961	GACCTGTGTTACTTAATCCGCTGCGCATCTCTCTCTCGTGGTCCGAGTGTGTGGGGTCTGTG	1020
Db	961	GACCTGTGTTACTTAATCCGCTGCGCATCTCTCTCTCGTGGTCCGAGTGTGTGGGGTCTGTG	1020
QY	1021	TGCGCAGCGATATCTGCTGCGGCACTGTGGGATCCATGAGCGGAGAGGGGCTGTGTCAGTGAATAC	1080
Db	1021	TGCGCAGCGATATCTGCTGCGGCACTGTGGGATCCATGAGCGGAGAGGGGCTGTGTCAGTGAATAC	1080
QY	1081	CGGCTGATAGCGTTGCGCTCGCGGGGGTAAACATGTTTCCCCACGCACTAATGTGCAAG	1140
Db	1081	CGGCTGATAGCGTTGCGCTCGCGGGGGTAAACATGTTTCCCCACGCACTAATGTGCAAG	1140
QY	1141	AGCGACGCGGAGCACTGTGTCATCTCAATCCCTTCGACTTACTATCAACCAACTGTGTG	1200
Db	1141	AGCGACGCGGAGCACTGTGTCATCTCAATCCCTTCGACTTACTATCAACCAACTGTGTG	1200
QY	1201	AAGAGGCTCCACAGTGAATTAACGAGGACTGTCCAGGCCGTGCTCGGCTCGTGAAG	1260
Db	1201	AAGAGGCTCCACAGTGAATTAACGAGGACTGTCCAGGCCGTGCTCGGCTCGTGAAG	1260
QY	1261	AGGGAATGTTTGGAGCTGGAATATGCAAGATTTTGGCTGAATTCAAGACTGTGCTCCAGTCC	1320
Db	1261	AGGGAATGTTTGGAGCTGGAATATGCAAGATTTTGGCTGAATTCAAGACTGTGCTCCAGTCC	1320
QY	1321	AAGCTCCGCGCGGCAATTAACGGGAGTCCCTTTTCTATGCAACGTTGGGTAAAGGGG	1380
Db	1321	AAGCTCCGCGCGGCAATTAACGGGAGTCCCTTTTCTATGCAACGTTGGGTAAAGGGG	1380

QY	1381	GTCTGCGGGGAGACGGCATCATGACAGACCACTGCTCATGTGGAGACAGATCAACGGGA	1440
Db	1381	GTCTGCGGGGAGACGGCATCATGACAGACCACTGCTCATGTGGAGACAGATCAACGGGA	1440
QY	1441	CATGTCAAAAAAGGTTCCATGAGGATCGTTGGGCTTAAGACTGTGTAGTAACATGTGGCAT	1500
Db	1441	CATGTCAAAAAAGGTTCCATGAGGATCGTTGGGCTTAAGACTGTGTAGTAACATGTGGCAT	1500
QY	1501	GGAACTATCCCATCAACAGCTAATACACACGGGGCCCCCTGCAACGCCCTCCCAAGGGCCAAAC	1566
Db	1501	GGAACTATCCCATCAACAGCTAATACACACGGGGCCCCCTGCAACGCCCTCCCAAGGGCCAAAC	1566
QY	1561	TATTCACAGGGGCGCTGTGGCGGGGTGGCTGCTGAGAGATGACGTGGAGTTACGGCGGGTGGGG	1620
Db	1561	TATTCACAGGGGCGCTGTGGCGGGGTGGCTGCTGAGAGATGACGTGGAGTTACGGCGGGTGGGG	1620
QY	1621	GATTTCCACTACGTGACGAGCATGACCACTGACCAAGCTAAATATGCCGTGCCAGGTTCCA	1680
Db	1621	GATTTCCACTACGTGACGAGCATGACCACTGACCAAGCTAAATATGCCGTGCCAGGTTCCA	1680
QY	1681	GCCCCCGCAATCTTCAACAGAAAGTGAGTGGGTGGCTGCAACAGTACAGCTCCGGCGGTGC	1740
Db	1681	GCCCCCGCAATCTTCAACAGAAAGTGAGTGGGTGGCTGCAACAGTACAGCTCCGGCGGTGC	1740
QY	1741	AAACCTCTCCCTACGGGAGAGAGTGCACATTCAGAGTCGGGCTCAACCAATACCTTGTTGGG	1800
Db	1741	AAACCTCTCCCTACGGGAGAGAGTGCACATTCAGAGTCGGGCTCAACCAATACCTTGTTGGG	1800
QY	1801	TCGCAGCTCCCATGCGAGCCCGAAACCGGATGTAGCAGTGCCTCACTTCATGCTCACCGAC	1860
Db	1801	TCGCAGCTCCCATGCGAGCCCGAAACCGGATGTAGCAGTGCCTCACTTCATGCTCACCGAC	1860
QY	1861	CCCTCCCACTCAACAGAGAGACGGCTAAACGACAGGTGGCCAGGGGGTCTCCCCCTCC	1920
Db	1861	CCCTCCCACTCAACAGAGAGACGGCTAAACGACAGGTGGCCAGGGGGTCTCCCCCTCC	1920
QY	1921	TTGGCCAGCTCTTTCAGCTAGCAGCAAGTTGTCTGCGCCTTCTCTGAAGGCGACATACATTACC	1980
Db	1921	TTGGCCAGCTCTTTCAGCTAGCAGCAAGTTGTCTGCGCCTTCTCTGAAGGCGACATACATTACC	1980
QY	1981	CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACTCTGTGGCGGATGAGATG	2040
Db	1981	CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACTCTGTGGCGGATGAGATG	2040
QY	2041	GGC 2043	
Db	2041	GGC 2043	
RESULT 6			
LOCUS	A76577	7065 bp	DNA
DEFINITION	Sequence 9 from Patent W09317110.	linear	PAT 19-OCT-1999
ACCESSION	A76577		
VERSION	A76577.1	GI:6088488	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 7065)		
AUTHORS	Rodgers,B.C. and Parker,D.		
TITLE	A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE		
JOURNAL	Patent: WO 931710-A 9 02-SEP-1993;		
FEATURES	WELLCOME FOUND (GB)		
source	Location/Qualifiers		
	1..7065		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
	<1..7065		
	/note="Encoded putative PT-NANBH non-structural proteins"		
	/codon_start=1		

Db 3604 GTCTGGCGGGAGACGGCATCATGACACCACTGCTCATGTGAGACACAGATCAACCGGA 3663
Qy 1441 CATGTCAAAACGCGTTCCATGAGGATCGTTGGGCGCTAAGACCTGTAGTAACATGTCGAT 1500
Db 3664 CATGTCAAAACGCGTTCCATGAGGATCGTTGGGCGCTAAGACCTGTAGTAACATGTCGAT 3723
Qy 1501 GGACATTTCCCATCAACGATACACACAGGCGCCCTGACGCGCTCCCGACGCGCAAC 1560
Db 3724 GGACATTTCCCATCAACGATACACACAGGCGCCCTGACGCGCTCCCGACGCGCAAC 3783
Qy 1561 TATTTCCAGGCGCGCTGTGGCGGGTGTGCTGTGAGAGATAGTGAAGTTACGCGGGTGGG 1620
Db 3784 TATTTCCAGGCGCGCTGTGGCGGGTGTGCTGTGAGAGATAGTGAAGTTACGCGGGTGGG 3843
Qy 1621 GATTTCCATGAGTACGAGCATGACCACTGCAACAGTAAATGCGCGCGGACAGGTTCA 1680
Db 3844 GATTTCCATGAGTACGAGCATGACCACTGCAACAGTAAATGCGCGCGGACAGGTTCA 3903
Qy 1681 GCGCCCGAATTTCTTCAACAGATGATGAGGCTGCGGCTGACAGAGTACGCTCCGCGTGC 1740
Db 3904 GCGCCCGAATTTCTTCAACAGATGATGAGGCTGCGGCTGACAGAGTACGCTCCGCGTGC 3963
Qy 1741 AAACCTCTCTTACGGGAGAGGTCATCTTCCAGTGGGCTCAACCAATACCTGTTGGG 1800
Db 3964 AAACCTCTCTTACGGGAGAGGTCATCTTCCAGTGGGCTCAACCAATACCTGTTGGG 4023
Qy 1801 TCGAGCTTCCATGAGGCGCGAACCAGATGAGAGTGTGAGTGTGCTTCAATGCTCACCGAC 1860
Db 4024 TCGAGCTTCCATGAGGCGCGAACCAGATGAGAGTGTGAGTGTGCTTCAATGCTCACCGAC 4083
Qy 1861 CCTCCCATCATCAGACAGAGAGCGGCTAAGCGCAGGCTGCGCAGGCGGCTCTCCCGCTCC 1920
Db 4084 CCTCCCATCATCAGACAGAGAGCGGCTAAGCGCAGGCTGCGCAGGCGGCTCTCCCGCTCC 4143
Qy 1921 TTGGCAGCTCTTACGAGTACGAGTGTGTGCGGCTTCTTCAAGGCGCATATACCTTAC 1980
Db 4144 TTGGCAGCTCTTACGAGTACGAGTGTGTGCGGCTTCTTCAAGGCGCATATACCTTAC 4203
Qy 1981 CAAATGACTTCCGAGAGCGTGAACCTCATGAGGCGCAACCTCTGTGGCGGATGAGATG 2040
Db 4204 CAAATGACTTCCGAGAGCGTGAACCTCATGAGGCGCAACCTCTGTGGCGGATGAGATG 4263
Qy 2041 GGC 2043
Db 4264 GGC 4266

RESULT 7
A32203 3750 bp DNA linear PART 10-DEC-1996
LOCUS NANBHV PT polypeptide fragment.
DEFINITION A32203
ACCESSION A32203
VERSION A32203.1 GI:1926564
KEYWORDS
SOURCE Non-A, non-B hepatitis virus
ORGANISM Non-A, non-B hepatitis virus
REFERENCE Viruses; unclassified viruses.
1 (bases 1 to 3750)
AUTHORS
JOURNAL Patent: FR 2655990-A 22 21-JUN-1991;
FEATURES Location/Qualifiers
source 1..3750
/organism="Non-A, non-B hepatitis virus"
/mol_type="unassigned DNA"
/db_xref="taxon:12440"
<1..>3750
/codon_start=1
/product="partial NANBH-PT polypeptide"
/protein_id="CAA02095.1"
/translation="MEGVFTGLTHVDHFLSQTQKAGNPFYLVAVQATVCARAQAP
PSMDQMKCLRLKLTGPTPLVRLGAVQNEVTLTHPTIFIAQMSADLEVVTST
WLVGVLAALAAVCLTGTGSVIVIRIILSGRAIVPREVLVGFDEMEBCASHLPY

IEQGMQLAEQFKQKALGLLQATYKQAEAAAPVESKRWALETFWAKHMMNTISG1QYL
AGLSTLPENPAIASIAIAPFASVTSEPLTLLNLIILGAVWPNSLPPSAAFAFGAG
IAGAAGSISGKQVLVDIAGAGVAAALVAFKMSGEMSEDEDLNMLIPALISAG
LVGVGAKAIIIAHVRGEGAVOMNNRILAPSRGNHVSPTHYVESAPAAARVQILS
DLITQULKRLHQHINEDCSTPCGSSWLRDWDWICTYLAQPKTMLOSKLPRLEGVP
FFSCORGKYGWRGDIMQTTCSGAQITGVHKNQSNMIVGPKTSSNMHGTFPIMAY
TTGCTPSPAPNYPALWRAEVEVETRGDFHYVTSMTDNKACQVAPAEFT
EVDGVRHLRVAIPACKPLREBVTQVGNQVLYVSGQLCEBEPDVAIVTMDPSHI
TAETKRLAIGSPSILASSASQLSASRATYITONDPRDADIIEANLIRHEMG
DITRVSENRKVILIDSFDPLAREDEREVSVPATIRKSKKPPMPAMPARDVNPPL
LESMAKDPYVPVHAGCLPPTKTPPIPPRRKRVUTESTVSALAEATKAGSS
EPSAVDSGTATAPDPQSDDGAGSDVSVSMPLBEPDGPLSDGSMSVSREAG
EDVCCSMSTYTMGTALITPCAAEBSKPIINALSNLHNMHNYATTSRCSAQRQK
TFEDRLQILDHYQDVLKEMKAKASTVKAKLISVEBACLTTPHSAKSKFYGAKDVRN
LSSKALNIHRISVMEDELTETPIDITITMAKNEVCCQPERGKRPALIVPELQVGR
VCEKRALVDVSTLPOAVMGSSYGVGRGROVERELVAMSKSTKTPMGAFTDTRCDS
VTENEDIVRESITVCCCLAPBAROAIISLRLRYIGPPLNRKQNGQYRRCRASGV
LITSQNLITCYLTKASAAACRAKLDQCTMLVCGDLVVICBSAGTQEDPAASLRVTEA
MTRVSAPBDPPEYDLELITSSNSNVAHDSGKRVYLLTRDP"

ORIGIN

Query Match 99.0%; Score 2022.2; DB 6; Length 3750;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2030; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 TGGAGGCGGCTTCTTCAACAGGCGCTTCAACGAGTGCAGCTTCTGCTCCCAACAAAG 60
Db 1 TGGAGGCGGCTTCTTCAACAGGCGCTTCAACGAGTGCAGCTTCTGCTCCCAACAAAG 60
Qy 61 CAGGAGGAGACAACTTCCCTTACTGTGTGAGGCTAAGCTGTGTGCGCTAGGCGC 120
Db 61 CAGGAGGAGACAACTTCCCTTACTGTGTGAGGCTAAGCTGTGTGCGCTAGGCGC 120
Qy 121 CAGGCGGCGCTTCTTCAACAGGCGCTTCAACGAGTGCAGCTTCTGCTCCCAACAAAG 180
Db 121 CAGGCGGCGCTTCTTCAACAGGCGCTTCAACGAGTGCAGCTTCTGCTCCCAACAAAG 180
Qy 181 CTGGCGGAGGCGCAACCTTGTGTATGAGGCTGAGGCGGCTTCAACGAGTGCAGCTTCT 240
Db 181 CTGGCGGAGGCGCAACCTTGTGTATGAGGCTGAGGCGGCTTCAACGAGTGCAGCTTCT 240
Qy 241 ACACACCCCATTAACCAATTCATCATGAGCATGATGATGATGATGATGATGATGATGATG 300
Db 241 ACACACCCCATTAACCAATTCATCATGAGCATGATGATGATGATGATGATGATGATGATG 300
Qy 301 AGCACCTGGGT 360
Db 301 AGCACCTGGGT 360
Qy 361 GGCAGCGGTGATGT 420
Db 361 GGCAGCGGTGATGT 420
Qy 421 AGGGAAGTCTCTTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Db 421 AGGGAAGTCTCTTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Qy 481 ATCGAGCGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 ATCGAGCGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy 541 ACAGCCACCAAGCAAGCGGAGGCGCTGCTCCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db 541 ACAGCCACCAAGCAAGCGGAGGCGCTGCTCCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy 601 GAGACCTTCTGGGCGAACAATGGAATCTTCAACGCGGATACAGTACTTACAGAGC 660
Db 601 GAGACCTTCTGGGCGAACAATGGAATCTTCAACGCGGATACAGTACTTACAGAGC 660
Qy 661 TTGTCACTCTGCTCGGGAATCCCGGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 TTGTCACTCTGCTCGGGAATCCCGGATGATGATGATGATGATGATGATGATGATGATGATG 720

OY	721	CTGTGCCCCGACACACCCAACTCAACCTCGCTGTAAACAATCCMGGGGGAGTGGATAGCC	780
Db	721	ACTGACCCTGTACACCAACCAATCTAACCTCTGTCTTAACTCTGGGGAGATGGGATCCG	780
OY	781	GCCCAACTCGCTCCCCCAGTGTGCTTTCAGCTTTCGTAGAGCGCGCATTTGCTGTGCG	840
Db	781	CCCAACTCGCTCCCCCAGTGTGCTTTCAGCTTTCGTAGAGCGCGCATTTGCTGTGCG	840
OY	841	GCTGTGGCAGCATAGGCTTTTGGGAAGGTGCTTGTGACATCTTGGCGGGCTTATGAGCA	900
Db	841	GCTGTGGCAGCATAGGCTTTTGGGAAGGTGCTTGTGACATCTTGGCGGGCTTATGAGCA	900
OY	901	GGAGTGGAGGCGCGCTGTGGGCTTTTAAAGTCAATGAGCGGCGAAATGCCCCCAACGGAG	960
Db	901	GGAGTGGAGGCGCGCTGTGGGCTTTTAAAGTCAATGAGCGGCGAAATGCCCCCAACGGAG	960
OY	961	GACCTGGTTAACTTACTCCCTGTGCCACTCTCTCTCGTGGTGCCTGATCGTGTGGGGTGTG	1020
Db	961	GACCTGGTTAACTTACTCCCTGTGCCACTCTCTCTCGTGGTGCCTGATCGTGTGGGGTGTG	1020
OY	1021	TGCGCAGCGATACTGCGTGGSCAGTGGGTCCAGGGAGGGGGGCTGTGCACTGATGAAC	1080
Db	1021	TGCGCAGCGATACTGCGTGGSCAGTGGGTCCAGGGAGGGGGGCTGTGCACTGATGAAC	1080
OY	1081	CGGCTGATAGAGCTTGCCTCGCGGGGTAAACAATGTTTCCCCCAGCACTAATGTGCCAGAG	1140
Db	1081	CGGCTGATAGAGCTTGCCTCGCGGGGTAAACAATGTTTCCCCCAGCACTAATGTGCCAGAG	1140
OY	1141	AGCGACGCCGAGCAGCTGTCACTCAGATCTCTCTCCGACTTACTATCAACCAACTGTG	1200
Db	1141	AGCGACGCCGAGCAGCTGTCACTCAGATCTCTCTCCGACTTACTATCAACCAACTGTG	1200
OY	1201	AAGAGGCTCCACCACTGATTTAACGAGGACTGCTTCCACGCGCTGCTCGAGCTGTGGCTA	1260
Db	1201	AAGAGGCTCCACCACTGATTTAACGAGGACTGCTTCCACGCGCTGCTCGAGCTGTGGCTA	1260
OY	1261	AGGATGTTTGGGACTGTGATATGCAACAATTTTGGGTGACTTCAAGACTGTGCTCCAGTCC	1320
Db	1261	AGGATGTTTGGGACTGTGATATGCAACAATTTTGGGTGACTTCAAGACTGTGCTCCAGTCC	1320
OY	1321	AAGCTCTCGCGCGCAATPACGGGGAGTCCCTTTTTCATATCCCAACGTGGGTACAAGGGG	1380
Db	1321	AAGCTCTCGCGCGCAATPACGGGGAGTCCCTTTTTCATATCCCAACGTGGGTACAAGGGG	1380
OY	1381	GTCTGGCGGGGAGACGGGATCATGACAGCAACTGTCTCATGTGAGACAGATCAACGGGA	1440
Db	1381	GTCTGGCGGGGAGACGGGATCATGACAGCAACTGTCTCATGTGAGACAGATCAACGGGA	1440
OY	1441	CATGTCAAAAACGGTTTCCATGAGAGTGTGGGCTTAAAGCTGTGATGAATGTGGCAT	1500
Db	1441	CATGTCAAAAACGGTTTCCATGAGAGTGTGGGCTTAAAGCTGTGATGAATGTGGCAT	1500
OY	1501	GGAACATTTCCCATCAAGCAATPACCAACGGGGCCCTGTACAGCCCTCCCAAGCGCAAAAC	1560
Db	1501	GGAACATTTCCCATCAAGCAATPACCAACGGGGCCCTGTACAGCCCTCCCAAGCGCAAAAC	1560
OY	1561	TATTCCAAGGCGCTGTGGCGGGTGTGCTGTGAAGAGTACGTGAGAGTTACCGCGGGTGGG	1620
Db	1561	TATTCCAAGGCGCTGTGGCGGGTGTGCTGTGAAGAGTACGTGAGAGTTACCGCGGGTGGG	1620
OY	1621	GATTTTCCACTACGTGACGAGCTGACCACTGCAACGTTAAATGCCCCGTGCAGAGTTTCA	1680
Db	1621	GATTTTCCACTACGTGACGAGCTGACCACTGCAACGTTAAATGCCCCGTGCAGAGTTTCA	1680
OY	1681	GCCCCCGAATTTTCAAGAAATGTGATGGGGTGGCGGTGCAACAGGTACGCTTCGGCGTGC	1740
Db	1681	GCCCCCGAATTTTCAAGAAATGTGATGGGGTGGCGGTGCAACAGGTACGCTTCGGCGTGC	1740
OY	1741	AAACCTCTCTACGGGAGAGGTGCATTTCCAGGTTCGGGCTCAACATACCTGTGTTGGG	1800
Db	1741	AAACCTCTCTACGGGAGAGGTGCATTTCCAGGTTCGGGCTCAACATACCTGTGTTGGG	1800
OY	1801	TGCAAGCTCCCATGCGAGCCCGAACCGGATGTAGAGTGTCACTTCCATGCTCAACCGAC	1860

Db	1801	TCGAGAGCTCCATGCGAGCCCGAAGCGAGTGTAGCACTGCTCACTTCAGTCTCAACCGAC	1860
Qy	1861	CCCTCCACATCACAGACAGACGGCTTAAGCGAGCTGCGCAGGGGGTCTCCCCCTCC	1920
Db	1861	CCCTCCACATCAACAGACAGACGGCTTAAGCGAGCTGCGCAGGGGGTCTCCCCCTCC	1920
Qy	1921	TTGGCCAGCTCTTACGTAGCAAGTTTGTCTGCGGCTTCCCTGAAAGCGGACATCATTAAC	1980
Db	1921	TTGGCCAGCTCTTACGTAGCAAGTTTGTCTGCGGCTTCCCTGAAAGCGGACATCATTAAC	1980
Qy	1981	CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACTCCTGTGGCGCATGAGATG	2040
Db	1981	CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACTCCTGTGGCGCATGAGATG	2040
Qy	2041	GGC 2043	
Db	2041	GGC 2043	
RESULT 8			
HPCK1S2		9410 bp	RNA
LOCUS			linear
DEFINITION			VR1 10-FEB-1999
ACCESSION	HPCK1S2	Hepatitis C virus (strain HCV-1b, clone HCV-K1-S2), complete genome	
VERSION	D50485		
KEYWORDS	D50485.1	GI:1030704	
	polyprotein; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A; NS5B; envelope protein; non-structural protein; interferon-sensitive; interferon-resistant; IFN-sensitive; IFN-resistant; ISDR; interferon sensitivity determining region; HVR; hypervariable region.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1	(sites)	
AUTHORS	Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Izumi, N., Marumo, F. and Sato, C.		
TITLE	Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region		
JOURNAL	J. Clin. Invest. 96 (1), 224-230 (1995)		
MEDLINE	95340824		
PUBMED	7542279		
REFERENCE	2	(bases 1 to 9410)	
AUTHORS	Enomoto, N.		
JOURNAL	Unpublished		
REFERENCE	3	(bases 1 to 9410)	
AUTHORS	Enomoto, N.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental University, Second Department of Internal Medicine, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113, Japan (E-mail: PXN04522@niftyserve.or.jp, Tel:03-3813-6111(ex.3224), Fax:03-3818-7177)		
FEATURES	Location/Qualifiers		
source	1..9410		
	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/strain="HCV-1b"		
	/db_xref="taxon:11103"		
	/clone="HCV-K1-S2"		
	/note="IFN-sensitive HCV from patient 2"		
	330..9362		
	/codon_start=1		
	/product="polyprotein"		
	/protein_id="BA09076.1"		
	/db_xref="GI:1814089"		
	/translation="MSTNPKPQRKTKNTRRNPQDYKPGGGQIVGGVILLPRRGPRLGVARTRKTSERSGRGRROPFKARQPEGMAQPGVPMPLYGNEGMAQWGLSPRGWLSRPMGPTDPRRSRRGKVIDLTGCFADLMGVIPLVGAIPLGASALAHGVLEDDGVNATGNTLPGCCSSTIFLALISLTLTPASAEVRNMSGVYHTNDCSNASIVYEADMTMTIPGCVPEVRANSSRCWALFTPLAANSSVPTTTRRHVDLLVGAAAFCSANT		
CDS			

VDLGSGVFLVSQLTFTSPRRHETVQDNCISYFGHVSGRHMANMNMNSPTTALVV
 SGLRLIPQAVMDMVGAAHGVLAGLAAYSMVNTAKNVLIVMLFPGVGRITVTCGAQ
 GHTTORLASFTFGAQRLOLVNTNGSMHINTALNCLNDLGLTGLVLAALFYHRNSS
 GCPBRILASCRSIDAFAGQSPITTAEPGSDORPYCMWAPRPGCIIVASRVCQVVC
 FPRSPVVGTTIDRSCAPTYSGENETDILLNTPRAPOGMMCTPMAGTCTKTCGG
 PPCNIGGVGNNTLCTPDCFRKHPKATYTKCSGPMWLPKRCVADPYRLMHPCTVNF
 TIFKRWYVGHVEHRLNAAICMWTGERCDLEDRDSELSPLLSTEMOOLICSPTTL
 PVLSTGLHLHONIVDOVLYGVSAVSVFKMEYIILLFLLDARCACTLMMMLL
 IAOAALLENVLVNAASVAGAHGFLSFLFCAWYIKGKLPVPAAGVAFVYVWELL
 LIALPRAVAMDRMAASCGAVFVGLALILSPHYVYVLAIRLWMOYFTTBAC
 LOWMPLPNAVGRAPAILITCALHPELIFSTIKLVAILGSLMTLOKSTMPFVR
 AOGILRACMLVRKVAAGHYVQAFMKLAALTGYVDHITPQDMAHGLMPLAIVE
 PVFSDMETKILITWADTAACGDIILGLASVARKKEIFLGPADLBESGMRLAAVE
 IAYSOQTRILGCIITSLTGRDQVEGEQVSTATQSLACTGTCVATVHGAGSK
 TLASGKPIITOMYTNVDOLVGMQAPSGARSILPCTGSSDLVLTGRADVYVVRG
 DGRSLSPRPVSVTKSSGGBLPCSHGAVGIPRAAOTRGAVALYIPRESMETT
 MRSPIYTNSSSPRVPOTFOVAHLHAPFGSGSKNKPVEYEAQCKVAVLINVSATL
 FGAVMSRAHGVDPNIRTGVRTITTGAPITYSTYKFLADGCGSGADYIILCDECHS
 TDSISLIGITVLDOAETAGARLVLAATAPGSAVTPHPNIEVALNIGBIPYK
 AIPETIKGRHLIFCHSKKCDLAAKSGGINAVYRGIDSVIPASGDVVVA
 TDLMTGTFGDPDSVIDNCTVOTVDSLDPTFIETTVPODASORSORGRG
 RAGIYRVPGRBPSGMDPSVLCQCYAGACAWVELTAPETSURLRAYINTGELVMO
 DHEFMSVFTGTLTHIDAFISQTKQAGNPPYLVAYQATVCAKQAAPPSMDQMKC
 LIRKPTLHGTPPLLYRUGAVQNEVILTHPIKYIMACWSADLEVSTWLVGVLA
 ALAACTTGSVIVGRIILSGKPAIIPDREVLRYEPMEECASHLPYIEGMLAE
 OFKOKALGILQTAQOAEAAAVESKMOALETFWAKMMNFIISGIVYLAGSTLPGN
 PALASIMFTASITSPITTOHTLFINIIGWAVAOALPPRAASAVFGAIAAAGSI
 CLGKTVNDIAGQAGVAGALVAFKMGSDMSTEDIVNLLPAILSPRATVGVCA
 LIRHVPBEGAVQMMNRLIAPASGNHSPHYVESDAARVQIISLITLOLK
 RLHOWINBDCSTPSCGAQITGHVKNKSMIVGPKCSNMHTGTPINAYTSCQKQK
 GVMRGDMQITPCGCAQITGHVKNKSMIVGPKCSNMHTGTPINAYTSCQKQK
 APNYSRALMRVAEVEYTRVGDHYVTGTDNLKCPQVAPAEFEJEDGVLHR
 YAPASKPLIDRETFQVGINOVVGSOLPCEPEPDVALVILNLTDSHTAARBL
 ARGSPSLASSASQSLAPSKATCTTRHSDPADVILANLIMRDEKNGITRYSEN
 KVLILDSFEPFLAEDEBESLPASILKSKRFPAMPPLWADPNPLIESKPDY
 VPPVHGPCLPPTKAPPIPPRRKRTVLTSTVSALAEATKTFGSESSAADSOT
 ATAPDQSDNDAGSDVESYSMPLBEPDPLSDGSMSTVEEAGEEDVCCMS
 YTMGTALITPCAEBSKLPINALSNSLRHNMVATTSRSASOKQKPTFRLDVL
 DHYRVLKEMKASTVAKRLVBEACKLTPHRSARSKFGVAGADVNLSTLNI
 RSVMKDLIEDTETPIDTTIMAKNEVFCQPEKGRKPARLIVPPLDGRVCKMALYD
 VSTLPQAVMGPSYGFQYSPQGRVEFLVNAWSKCKPWFPAIDTCEPSTVESIRV
 BESIYQCDLAPEARQARSLTERLYIGPLTNSGONGCYGRCSAGVLTJESNITL
 TCYLKATACRAKAKQDCITMLVAGCDLVYIICESAGSTVETEMETYSAPG
 DPQOYVLELITSCSNVSAHDSAGRVYVLTDPPTPLAARAMEYARHPVNSWL
 GNIMVAPTLAARMLIMTHFESILLAGOLLEKALDQCYGACYSIEPDIPTIOLRH
 GLSARSLHSYSGEINRVAASCLKRGVPLRWRHRBASVRKLLISGGRALICGRT
 FNMVATRLKLTPIPAASQLDLSGMFVAGSGDIYHSVSRAPRPMFWCILLLSVG
 GIYLPNR
 330. .902
 /product="core protein"
 903. .1478
 /product="E1"
 1479. .2756
 /product="E2"
 2757. .3407
 /product="NS2"
 3408. .5300
 /product="NS3"
 5301. .5462
 /product="NS4A"
 5463. .6245
 /product="NS4B"
 6246. .7586
 /product="NS5A"
 7587. .9359
 /product="NS5B"
 1479. .1559
 /note="hypervariable region (HVR) "
 6954. .7073
 /note="interferon sensitivity determining region (ISDR) "
 misc_feature
 misc_feature
 Query Match

88.3%; Score 1804.6; DB 14; Length 9410;

Best Local Similarity 92.7%; Pred. No. 0;
 Matches 1894; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
 1 TGGGAGGGGCTTTTCAAGGCTCAACCAAGGATGACCACTTCCTGACCAACAAAG 60
 5001 TGGGAGGGGCTTTTCAAGGCTCAACCAAGGATGACCACTTCCTGACCAACTAG 5060
 61 CAGGAGAGAGCAACTTCCCTCACTGCTGAGGATGACCAAGCTACTGTGCGCTAGGGC 120
 5061 CAGGAGAGAGCAACTTCCCTCACTGCTGAGGATGACCAAGCTACTGTGCGCGCTAG 5120
 121 CAGGCTCACTTCATCATGAGGATGAAATGTGGAAGTGTCTCATACGGCTTAAGCTTACT 180
 5121 CAGGCTCACTTCATCATGAGGATGAAATGTGGAAGTGTCTCATACGGCTTAAGCTTACG 5180
 181 CTGGCGGGGGCAACCCCTGCTGATATGAGCTGGGAGCGGTCAAAAGAGGTCAACCTC 240
 5181 CTGGCGGGGGCAACCCCTGCTGATATGAGCTGGGAGCGGTCAAAAGAGGTCAATCTTC 5240
 241 ACACACCCCATTAACCAATTCAATCATGAGCATGATGACCGACCTGAGAGTCTGACG 300
 5241 ACACACCCCATTAACCAATTCAATCATGAGCATGATGACCGACCTGAGAGTCTGACG 5300
 301 AGCACCCTGGGTCTGCTGAGGCGGGGTCTTGACGCTCTGCTGCTGCTTATGCTTGAACA 360
 5301 AGCACCCTGGGTCTGCTGAGGCGGGGTCTTGACGCTCTGCTGCTGCTTATGCTTGAACA 5360
 361 GCGACCGGTGATATGCTGAGTATGATCATCTTGTCCGGCGCGGCTATGCTTCCGAC 420
 5361 GCGACCGGTGATATGCTGAGTATGATCATCTTGTCCGGCGCGGCTATGCTTCCGAC 5420
 421 AGGGAAGTCTCTTCAAGAGATTCATGATGATGAGAGTGGCGCTGCACTCCCTTAC 480
 5421 AGGGAAGTCTCTTCAAGAGATTCATGATGATGAGAGTGGCGCTGCACTCCCTTAC 5480
 481 ATGAGCAAGGAGTACGCTGCGGAGATTCAGCAAAAGCGCTGGGTGCTGACG 540
 5481 ATGAGCAAGGAGTACGCTGCGGAGATTCAGCAAAAGCGCTGGGTGCTGACG 5540
 541 ACGACCAACCAAGCAAGCGGAGCGGCTGCTCCCGTGTGTGAGTCCAAATGGGAGACCTT 600
 5541 ACGACCAACCAAGCAAGCGGAGCGGCTGCTCCCGTGTGTGAGTCCAAATGGGAGACCTT 5600
 601 GAGACCTTCTGGGCGAAACCATGTGGAATTCATCAGCGGATCAATATCTTACAGGC 660
 5601 GAGACCTTCTGGGCGAAACCATGTGGAATTCATCAGCGGATCAATATCTTACAGGC 5660
 661 TTGTCCACTCTGCTGGGAATCCCGATGTCATCATGATGCGCTTCAACAGCTCTGTCT 720
 5661 TTGTCCACTCTGCTGGGAATCCCGATGTCATCATGATGCGCTTCAACAGCTCTGTCT 5720
 721 ACTAGCCGCTCAACCAATCTACCTCTGCTTAACTCTCTGGGGAGATGGTATGCC 780
 5721 ACTAGCCGCTCAACCAATCTACCTCTGCTTAACTCTCTGGGGAGATGGTATGGTGC 5780
 781 GCCCACTGCTGCTCCCGAGATGCTGCTTCAAGCTTCTGATGAGCGCTGCGATGCG 840
 5781 GCCCACTGCTGCTCCCGAGATGCTGCTTCAAGCTTCTGATGAGCGCTGCGATGCG 5840
 841 GCTGTGGCAGCATAGGCTTGGGAAGTGTGTGACATCTTGGCGGCTATGAGAGA 900
 5841 GCTGTGGCAGCATAGGCTTGGGAAGTGTGTGACATCTTGGCGGCTATGAGAGA 5900
 901 GGAAGTGGCAGGCGGCTGCTGCTTAAAGTGCATGAGCGCGGAAATGCTCTCCACGAG 960
 5901 GGAAGTGGCAGGCGGCTGCTGCTTAAAGTGCATGAGCGCGGAAATGCTCTCCACGAG 5960
 961 GACCTGTTAATCTACTCTGCTGCAATCTCTCTGCTGAGCCCTGCTGCTGCTGCTG 1020
 5961 GACCTGTTAATCTACTCTGCTGCAATCTCTCTGCTGAGCCCTGCTGCTGCTGCTG 6020
 1021 TGGCAGGAGTACTGCTGCTGCGACGCTGAGCTGCGAGGAGGAGGCTGTGCACTGATGAG 1080

Db 6021 TCGCAGCAATATCTGCGTCGCGATGTGGGCCCCGGAGAGGGGGCTGTGCATGTGATGAAC 6080
Qy 1081 CGCGTGAATAGCGTTGCGCTCGCGGGGTAAACATGTTTTCCCCCGACGACATATGTGCAGAG 1140
Db 6081 CGGCTGATAGCGTTGCGCTTCGCGGGGTAAACAGTCTCCCGACGACATATGTGCTGAG 6140
Qy 1141 AGCGAGCGCGACAGCTGTCACTCATGATCTCTCTCCGACTTACTATCAACCACTGTTG 1200
Db 6141 AGCGAGCGCGACAGCTGTCACTCATGATCTCTCTCCGACTTACTATCACTGAGCTGTTG 6200
Qy 1201 AAGAGGCTCCACAGTGAATTACAGAGAGCTGCTCCAGCGCCCTGCTCCGGGCTCGGGGCTA 1260
Db 6201 AAGAGGCTCCACAGTGAATTACAGAGAGCTGCTCCAGCGCATGCTCCGGCTCGGGGCTA 6260
Qy 1261 AGGAGTGTGGGAGCTGATATGACAGTTTTTGGCTGACTTTCAGACCTGCTGCTCAAGTCC 1320
Db 6261 AGGAGTGTGGGAGCTGATATGACAGCGTGTGACTGACTTCAAGCTTGGCTCCAAATCC 6320
Qy 1321 AAGCTTCCTGCGCGATTAACCGGGAGTCCCTTTTCTCATGCGCAACGTGGGTACAAAGGG 1380
Db 6321 AAGGTCCTGCGCGGTAAACAGAGTCCCTTCTCTCATGTCAACGTGGGTACAAAGGGA 6380
Qy 1381 GCTTCGCGGGGAGAGGAGCATCATGAGACCACTGCTCATGTGAGAGCAGATCAACCGGA 1440
Db 6381 GCTTCGCGGGGAGAGGAGCATCATGAGACCACTTGCCTCATGTGAGAGCAGATCAACCGGA 6440
Qy 1441 CATGTCAAAAACGTTTCCATGAGATCGTTGGGGCTTAAGACCTGTATGATTAACATGTGGCAT 1500
Db 6441 CATGTCAAAAACGTTTCCATGAGATCGTTGGGGCTTAAGACCTGTGAGCAATATGTGGCAT 6500
Qy 1501 GGAACATTTCCCATCAACGAGATCAACACGAGGAGCCCTGACAGCCCTCCCGACGCCAAGC 1560
Db 6501 GGAACATTTCCCATCAACGAGATCAACACGAGGAGCCCTGACAGCCCTCCCGACGCCAAGC 6560
Qy 1561 TATTTCAGGGGCGCTGCGGGGTGCTGAGAGATGATGAGAGTTTACGCGGGGTGGGG 1620
Db 6561 TATTTCAGGGGCGCTGCGGGGTGCTGAGAGATGATGAGAGTTTACGCGGGGTGGGG 6620
Qy 1621 GATTTTCACTACGATGACGAGATGACCACTGACCAACGTAAATGCGCCGTGCGAGGTTCCA 1680
Db 6621 GATTTTCACTACGATGACGAGGATGACCACTGACCAACGTAAATGCGCCGTGCGAGGTTCCCT 6680
Qy 1681 GCGCCCGAATTTTTCACAGAGTGGAGTGGCGCTGACACAGTACCGCTTCGCGCTGC 1740
Db 6681 GCGCCCGAATTTTTCACAGAGTGGAGTGGCGCTGACACAGTACCGCTTCGCGCTGC 6740
Qy 1741 AAACCTCTCTACGAGGAGGATGACATTCACAGTTCGCGGCTCAACCAATACCTGATGGG 1800
Db 6741 AAACCTCTCTACGAGGAGGATGAGTCAATTCACAGTTCGCGGCTCAACCAATACCTGATGGT 6800
Qy 1801 TCGCAGCTCCATGCGAGCGCCGACCGAGTGTAGAGTGTCTGCTCATGTCTCACCGAC 1860
Db 6801 TCGCAGCTCCATGCGAGCGCCGACCGAGTGTAGAGTGTCTGCTCATGTCTCACCGAC 6860
Qy 1861 CCCTCCCATATCAACAGCAGAGACCGGTAAAGCGAGGCTGAGCGAGGGGTCTTCCCTCC 1920
Db 6861 CCCTCCCATATCAACAGCAGAGCGCGGTAAAGCGGTAGGCTGAGCGAGGGGTCTTCCCTCC 6920
Qy 1921 TTGGCAGGCTCTTCAGCTGACGATGCTGTGGGCTTCTCTGAAAGCGGACATACATTAACC 1980
Db 6921 TTGGCAGGCTCTTCAGCTGACGATGCTGTGGGCTTCTCTGAAAGCGGACATACATTAACC 6980
Qy 1981 CAAATGACTTCCACAGACGCTGACCTCATGAGGCGCAACCTCTGTGGCGGATGAGATG 2040
Db 6981 CGTCAAGACTCCCGCAGATGCCAATCTCATGAGGCGCAACCTCTGTGGCGGATGAGATG 7040
Qy 2041 GGC 2043
Db 7041 GGC 7043

RESULT 9
AB049088

LOCUS AB049088 9616 bp RNA linear VRL 22-AUG-2002
DEFINITION Hepatitis C virus genomic RNA, complete genome, isolate:HCVT094.
ACCESSION AB049088
VERSION AB049088.1 GI:11559442
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
AUTHORS
REFERENCE
JOURNAL
TITLE
AUTHORS
FEATURES
source
1. .9616
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="HCVT094"
/specific_host="human"
/db_xref="taxon:11103"
/tissue_type="serum"
/note="HCV-genotype: 1b"
1. .341
/note="complete"
342..9374
/codon_start=1
/product="polyprotein"
/protein_id="BAB18801.1"
/db_xref="GI:11559443"
/translation="MSTNPKPQRTKRNTRNRPDVKFPGGQIGVGVYLLPRGRPL
GVRAKTRSERSGRGRROPILKARPERGRWADQVWPILYGNMGMAALMSPRG
SRPQKPTDPRRRSRNKGVIDITLTCPKADMGITPLVGAFLGGAAPLACGVULSD
SVNATGNTLPGCSFSILALISCTLPASVEVHNSGIHVNDGNSATVYBAD
MIMAPAGVPCPRDNDASRCWVALPTLAAENASVPTTIRHVDLVGAAPCSAMY
VGDLCGAVFLISQFTGSPRHEVVDNCSIIYGHVGHMADMMNMPALAV
SOLARIPQAVMDVAGAVGLAGLAAYSNMNAKYLIVLLFAGVDGNTHGAVQ
GRVYSLSANFISLASOKIQIINTNGSMHINRTALNCDLSLOTFLALFAHKPNAS
GCPBRMASCRSIDKFDGMPITYAOPSNDQRYCHRYAQQGIVPASQVCPYVC
FTSPVAVGTTDSSGVPPTYSGNENETDVLNLNTRPQGNFPGCTMANSIGTKTCGG
PPCNIGVGNNTLTCPTDCPKRHEBATYTKGSGPWLPRCLVYPRMLMYPCVNF
TIPIKRMVGVGEHRHLHAACMTGERPDLDRDRSELSPLLSITTEWQIIPCSFTTL
PALSTGLJHONIVDOYVLIGISAVSFAPIKMEYVLLFLLLADARVCACLMMMLL
IAQDAALBNIVYINMAASVAGAHGFLSGTPCAAWYLIKGRVLGVAVYALYGWILL
LLALPPRAYAMDREMASCGAVFVGLALITLTPYKVLARLIMWLQIYIAAEAH
LQWVPPPLNVRGGRDAIILLTCAHPELIFDITKLALILGLPWFQGIIRVYFVR
AQGILRYCMLVRKAVAGHYVQMALMKLALGTIVYHLPLRDMANGLDLVAVE
PVPFSDMETKILTWGADPAACGDIISGLPVARSGKEILLPADSLBGRGRLLAPIT
AYSOTGELGCIITSLGRKDNKEGVOVSTATOSFLASCNVGCVMTVYHAGSK
TLAQRKEIITOMYNNODLYGNKAPBGRBMPCTCGSSDLYLVTHHADIVPVRRG
DGRSLSPREVSYLKGSSGGLCPGSHVVGITRAAVCTRGAKANDPVESEMETT
MRSPVFTDNSSPPAVPQTFYAHLAHPGSGSKTKVPAAYAAQGXVLVLPVSVALT
GFGAYMSKAHSTDNIRGTITITGAPITVSTYGGFLADGCGGAGVDIICDCHS
TDSITLIIGITVLPQAEAGARLVLVLAATPSPGVTPHPNIEVALSNTGEIPYVGK
AIPLEAIKGRHLIFCSKKRCDRLAALSGGINALAYRGLDVSVPISGDDVVVA
TDALMTGFTGPDVIDNCTVYQTVDPESLDPTTIEITTYTPDAVASTORGRGRG
RGLIYREVTPEGRSPSGMDSVLCEDDAGCAMELFAETSVLRATLPPGDPVCO
DHLBFWESVFTGLTIDAHPLSQTKQADNCPYLVAYOATVCARQAPSPMDQWKC
LIRLKPTLHGPTPLLYRGAVNEVTLHTPTKXYIMCMSADLEAVTSTWLVGVGLA
ALAAVCLTGGSVIVGRIILSGRAVILIDREAVILREPEMECSAHLIYIQGMLAE
OFKOKALIGTATTOAALAAAPVBSKRALEAFQAKMMNPISGIVYLAGISTLPGN
PAISLMAFTASITSPULTOSTLFLNLTGAGLAPAPASAPVAGGAGTGAATGSI
GLGRVLVDILAGYAGVALVAFKWSGEMFSTEDLVNLPLAISLPAALVVGVC


```

Db      6693  ACCCCGAATCTTCAACGAGGTGATGAGTGCAGCGCTGCACAGGTACGCTCCGCGCTGC 6752
Qy      1741  AAACCTTCCTTACGAGGAGAGGTCAATTCAGAGTCGAGGCTCAACCAATACCTGGTGGG 1800
Db      6793  AAACCTTCCTTACGAGGAGAGGTGCTATTCAGAGTCGAGGCTCAACCAATATCTGGTGGG 6812
Qy      1801  TCGCAGCTCCATGCGAGGCGCCGACCGAGTGTAGCAGTGTCTCACTTCCATGCTCAAC 1860
Db      6813  TCACAGCTCCATGAGAGCGCCGACCGAGTGTAGCAGTGTCTCACTTCCATGCTCAAC 6872
Qy      1861  CCCTCCCATATACAGCAGAGCGGCTTAAGGCGAGGCTGAGGAGGCTTCCCCCTTCC 1920
Db      6873  CCCTCCCATATACAGCAGAGCGGCTTAAGGCGAGGCTGAGGAGGCTTCCCCCTTCC 6932
Qy      1921  TTGGCAGCTCTTCAAGTGTCTGAGGCTTCTCTCGAGGCGCATATACCTTACC 1980
Db      6933  TTGGCAGCTCTTCAAGTGTCTGAGGCTTCTCTCGAGGCGCATATACCTTACC 6992
Qy      1981  CAAATGACTTCCAGAGCGCTGACCTCATCGAGGCGCAACCTCTGTCGCGCATGAGATG 2040
Db      6993  CATGATGACTTCCGAGCGCTGACCTCATCGAGGCGCAACCTCTGTCGCGCATGAGATG 7052
Qy      2041  GGC 2043
Db      7053  GGC 7055

RESULT 10
AF333324 9587 bp mRNA linear VRL 18-DEC-2002
LOCUS AF333324 Hepatitis C virus type 1b polyprotein mRNA, complete cds.
DEFINITION AF333324
ACCESSION AF333324
VERSION AF333324.1 GI:12631192
KEYWORDS Hepatitis C virus type 1b
SOURCE Hepatitis C virus type 1b
ORGANISM Hepatitis C virus type 1b
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnaviruses.
REFERENCE 1 (bases 1 to 9587)
AUTHORS Thomson, M., Nascimbeni, M., Gonzales, S., Murthy, K.K., Rehmann, B.
and Liang, J.J.
TITLE Emergence of a distinct pattern of viral mutations in chimpanzees
infected with a homogeneous inoculum of hepatitis C virus
JOURNAL Gastroenterology 121 (5), 1226-1233 (2001)
PUBMED 21534507
REFERENCE 2 (bases 1 to 9587)
AUTHORS Thomson, M. and Liang, J.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2001) LDS, NIDDK/NIH, 10 Center Drive, Bethesda,
MD 20892, USA

FEATURES
source location/Qualifiers
1..9587
/organism="Hepatitis C virus type 1b"
/mol_type="mRNA"
/strain="chimera of HCV-BK, HCV-J, HCV-H"
/serotype="1b"
/specific_host="chimpanzee"
/db_xref="taxon:31647"
/clone="HCV-CG1b"
1..341
1..271
misc_feature
1..271
/notes="derived from BK strain; similar to sequence
deposited in GenBank Accession Number M58335"
272..3030
/notes="derived from J strain; similar to sequence
deposited in GenBank Accession Number D90208"
342..9374
/codon_start=1
/product="polyprotein"
/protein_id="AAK08509.1"
/db_xref="GI:12831193"
CDS

```

```

/translaction="MSTPQRKTKRNTNRPDVYKFGGQYVGVYLLIPRGRPL
GVRATRTKTSERQRRRQPIPKARPEGRTWAPGIPWLYGNGMGMALMSLRG
SRPMGPDTPRRSRRLKXVIDTLCEGADLMQYIPIVAGELGAGARLAGVLEL
GVNATYGLPCGSPSIFLALLSLCTIPASAYEVRANYSIHVINDGNSIIVEAD
MIMHGPVCPRESNSRSCWALPTLALPTLASISLPTTIRHVDLGAALGSLM
VGDLCGVYELVSLPTSPRRYETVODCNISYVGHSHGMADMNMNSPTTALVY
SGLRITPOAVVDMAKHWGLAGLAYSIVNGMAKULIYNLLRAGVDGHTHVGAV
ASSTOSVLSLQSPSOKILOVNTNGSHIRKTNALNCDSIQTGIALFTAHFNMS
GCPERRMASCRIIDKFAQGMPIITHVNPISQURCMYAPQPGIYASQVCEVC
FTSPVVGTTDRSGVPTYSWGENETDILLNTRPPGSMFGCTNMNSTFTKCTNF
PCNIGGVNNTLICPTDCFRKHPEATYTGSGSGLMPLPRCLVDYPIVLMYPCINF
TIPRYVAGVEHRIINAACWTRGERRDDEDRSELSPILLSTTEMQVPCSTTL
PALSTGLHLMQNVVDYOYLYGVGSYVSVYIKMEYITLLFLLADAPVCAIIMML
IAQEBATLENVINAASVAGAGLSTLFFCAWYIKGLVLEGAALYAGWPLML
LILALPRVAMDEREMASCAGAVFGLVLLTLPKYVFLARLIMLQYITREAH
LOVWPPLNLYGGRDAILLILCAHPELIPITLXLLALPLMWLOQIGIRVFPVR
AOGHILACMLVRKAGHYVOMAFMKGLAGTYIYNLTPRLBMAHAGLIDVAVE
PVPSDEETKXITMGADPAACGDIILGIPVARGKELILGADSLBERGRLAPIT
AYSQTRRLGCIITTLGRDRNOYBEGVYVSTRATCGSLQLQTVNHCWYVYHAGSK
TLAPKGEITQMYTNVDQDLVGMQAPBARSILPCTGSSDLQTVTRHADYIPRRRG
DSRGSLSPREVSYLKSSGGPLCPSSHAAGIIPRAACTGAKAAVDVPEVESMETT
MRSVFTDNSPPAVPOSFQYAHLAHAPGSGSKTKVPAAYAOQKXVLVNLPSVATL
GEGAYMSKAHGIDNIRGTGRTITGAVVTYSTYKFLADGCGGAGVDIILICEHS
TDSPTIIGITVLDQAEAGARLVILATPFGSVTPHPNIEVALSNTGEIPEYK
ALPIEATRGGRHLIFCHSKKCDLALSLGAINAVAYRGDVSYPITGDIVVVA
TDALMTGYTGDPSVDNCTVOTQVDPFSLDPTITLTTPYODANSLSQRRTRGRG
RGRGYREVTPEGRSGMFDSSVLCEDYDAGAWELPAETSVRLAYALNTPGLVCQ
DLHEFMEVSFTGLTHIDAHFLSOTKQADNPEYVAYAOATCARAOAPSPMDQMKC
LIRKPTLHGPPLTYRGLAVONETVTLTHDPTIKYIMACMSADLEVSTVTLVGVLA
ALAAVCLTTSQSVIVGRTIISGRPAIVDRLELLOEPMEECATVLIIEQMLAE
OFKQALGILDTATKQAEALAPVYBSKRALETFRAMAMNFIIGIOTVLAETSLP
PAISLMAFVASTISPLTOSTLLFNILGAVNAQALAPASAPAVGAGIAGAVSI
GLGAVLVDILAGYAGAGALVARKVMSGEMSTEDLVNLLPALISPALVGVCAA
LIRRHVGEAGVOMNRLIAPASRGHNSPTHVPEBDAARVATQILSLITQOLK
RLHOWINDECTPCSGSLRDYMDMICVLDPFTKLOSLLPOLPVPFSCORGYK
GWRMGDGMOTTCPCGAGIOTNGHVNKSGRWITGPTKSTMTGTFRINATYTGCTPSP
APNYSRALMRVAEIEYEVTVGDPHYTGTGTTDNVKKPCQVAPABEFSEVDGRHLR
YAPACRPLIREVTFVQVNLQYLVGSQLCEPEVAVLITMLDPSHTIATBARL
ARGSPPLASASASQLSAPLSKATCTTHVBPDDLILANLMLQDEMGNTITRVSEN
KVVLVDSFDPPLRAEDEREVSIPALIKSKSPFAPAMPIMARPDPNPLLSMXDPDY
VPPVHGCEPLPIKAPPIPPRRKRTVLTSSYSALAEIATKPTGSESSASVAGST
ATAPLPDQSDGDDGSDYESMSMPLGEGGDDPLSDGMSSTVSEBASBPVCCMS
YTMGALITTPCAABESKIPINALNSLIRHNMVYATTSRSAGLRQKRYTDRLOVD
DHYVDVLEKEMAKASTYVAKLISVEACKLIPSPHASKFEGAGDVNLSKAVNHI
HSWMLDLIEDTPIIDTITMAKNEVFCQPKRGKRPRLVLPFGVRCVAKMALYV
VSTLPQVMSVSGFOYSPGQVFEVLTVMKSKNPMGFSYDTECPSTYENDIRV
BESYOCDDLAPEARQATKSLTERLYIGGPLTNSKQNGRCYRCSASGLVLTSCNTL
TCYLKASAAACAAKQCTMLVNDPDLVTCESGQDDEASLASIRVFTAMRYSPG
DPQPEYDLELITSCSSVSVYAHDSGRVYITRTDPTPLARAWETARHTPVNWSL
GNIMYAPTLMARMLTMHFPISLLAQBOLKALDQCYGACYSIEPDLIPIOTERLH
GLSAFSLHSYSGEINRVASCIRKGVPLVWRHRRARSVBARILSOGGRATYCKYL
FNMAVTKLTKLTPIPASQDLDSGFVAGSGDYIHSLSRARPRFMILCLLLSVGV
GYLLPNR"
1..9404
/notes="derived from BK strain; similar to sequence
deposited in GenBank Accession Number M58335"
9375..9587
misc_feature
9405..9587
/notes="derived from HCV-1a strain H; similar to sequence
deposited in GenBank Accession Number AF009070"

ORIGIN
Query Match 88.2%; Score 1801.4; DB 14; Length 9587;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
Qy      1 TGGAGGGCGCTTTCACAGGCTTACCCAGTGTGCTTCTGCTCCCAACAAAG 60
Db      5013 TGGAGAGTGTCTTCAACAGGCTTCAACCAATAGATGACACTTGTCCACAGACAG 5072
Qy      61 CAGGAGAGACAACTTCCCTTACTGAGGCTGACAGGCTACGTGTGGCTAGGGCC 120
Db      5073 CAGGAGAGACAACTTCCCTTACTGAGTATACCAAGCAAGGTGTGCGCGGCT 5132

```

QY 121 CAGGCCCACTCATCATGAGATCAATGTGGAATGTCTCATAGGCTAAAGCTACT 180
 DB 5133 CAGGCCCACTCATCATGAGATCAATGTGGAATGTCTCATAGGCTAAAGCTACTAG 5192
 QY 181 CTGCGCGGCGCAACCCCTTGTGTATAGAGCTGGAGCCGTCGCAAAAGAGTCACTTC 240
 DB 5193 CTGCAACGGGCGCAACCCCTTGTGTATAGAGCTGGAGCCGTCGCAAAAGAGTCACTTC 5252
 QY 241 ACACACCCCATTAACCAATTCATCATGAGATGATCATGAGCCGACTGGAGGTCTGACAG 300
 DB 5253 ACCCAACCCCATTAACCAATTCATCATGAGATGATCATGAGCCGACTGGAGGTCTGACAG 5312
 QY 301 AGCACCCTGAGTCTGAGTGGCGGAGCTCTTGACAGCTGAGCTGCTGATTTGCTTGACACA 360
 DB 5313 AGCACCCTGAGTCTGAGTGGCGGAGCTCTTGACAGCTGAGCTGCTGATTTGCTTGACACA 5372
 QY 361 GGCAGCGTGTATGTGTGGGTAGATCATCTGTCCGGCGGCGCTATTTGTTCCGAC 420
 DB 5373 GGCAGTGTGTATGTGTGGGTAGATCATCTGTCCGGCGGCGCTATTTGTTCCGAC 5432
 QY 421 AGGGAATCTTACCAAGAGTTCCATGAGATGGAAGATGGCGCTGCACTCCCTTAC 480
 DB 5433 AGGGAATCTTACCAAGAGTTCCATGAGATGGAAGATGGCGCTGCACTCCCTTAC 5492
 QY 481 ATCGACAGGAGATGAGCTGCGCGGAGAGTTCAAGCAAAAGCGCTGGGTTGCTGAC 540
 DB 5493 ATCGACAGGAGATGAGCTGCGCGGAGAGTTCAAGCAAAAGCGCTGGGTTGCTGAC 5552
 QY 541 ACAGCCACCAAGCAAGCGGAGGCGCTGCTCCGCTGGTGGAGTCCAAAGTGGCGAGCCCT 600
 DB 5553 ACAGCCACCAAGCAAGCGGAGGCGCTGCTCCGCTGGTGGAGTCCAAAGTGGCGAGCCCT 5612
 QY 601 GAGACCTTGTGGGCGCAACATGTGAACTTTCATATAGCGGATACAGTACTTAAAGAGC 660
 DB 5613 GAGACCTTGTGGGCGCAACATGTGAACTTTCATATAGCGGATACAGTACTTAAAGAGC 5672
 QY 661 TTGTCACTCTGCTGGGAAATCCCGGATGTGATCATGATGGCGTTCAAGCTCTGTC 720
 DB 5673 TTGTCACTCTGCTGGGAAATCCCGGATGTGATCATGATGGCGTTCAAGCTCTGTC 5732
 QY 721 ACTAGCCGCTCAGCAACCCATCTACCTCTGCTTAAATCATCTGGGCGGATGGTAGCC 780
 DB 5733 ACTAGCCGCTCAGCAACCCATCTACCTCTGCTTAAATCATCTGGGCGGATGGTAGCC 5792
 QY 781 GCCCACTGCTCCCCAGAGTGTCTTCAAGCTTTGTAAGGCGCGGCAATTGCTGTGCG 840
 DB 5793 GCCCACTGCTCCCCAGAGTGTCTTCAAGCTTTGTAAGGCGCGGCAATTGCTGTGCG 5852
 QY 841 GCTGTGGCAGATAGGCTTGGGAAAGTGTGTGGAATCATCTGGCGGGCTATGAGAGA 900
 DB 5853 GCTGTGGCAGATAGGCTTGGGAAAGTGTGTGGAATCATCTGGCGGGCTATGAGAGA 5912
 QY 901 GGAAGTGGAGCGCGCTGCTGAGCTTTAAGTATAGGCGGCAAAATGCTCTCAACGAG 960
 DB 5913 GGAAGTGGAGCGCGCTGCTGAGCTTTAAGTATAGGCGGCAAAATGCTCTCAACGAG 5972
 QY 961 GACCTGTGTAATTAATCTCTGTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 DB 5973 GACCTGTGTAATTAATCTCTGTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 6032
 QY 1021 TGGCAGGAGATCTGCTGCTGAGACGAGGATTCAGAGGAGAGGAGCTGTGAGTGAAGAAC 1080
 DB 6033 TGGCAGGAGATCTGCTGCTGAGACGAGGATTCAGAGGAGAGGAGCTGTGAGTGAAGAAC 6092
 QY 1081 CGGCTGATAGCTTGCCTCGCGGGGTAAACATGTTTCCCAACGCACTATGTGCCAGAG 1140
 DB 6093 CGGCTGATAGCTTGCCTCGCGGGGTAAACATGTTTCCCAACGCACTATGTGCCAGAG 6152
 QY 1141 AGCGAGCGCGAGCAGTGTCACTCAGATCTCTCTGCACTTATCTACCCCACTGTTG 1200
 DB 6153 AGCGAGCGCGAGCAGTGTGTACTCAGATCTCTCTGCACTTATCTACCCCACTGCTGCTG 6212

QY 1201 AAGAGCTCCACAGTGAATTAACGAGAGCTGCTCCAGCCCTGCTCGGCTGTGCTA 1260
 DB 6213 AAGAGCTCCACAGTGAATTAATGAGAGCTGCTCCAGCCCGTGTCCGGCTGTGCTA 6272
 QY 1261 AAGAGTGTGGAGCTGATTAATGACAGTGTGGCTGACTTCAAGACCTGCTCCAGTCC 1320
 DB 6273 AAGAGTGTGGAGCTGATTAATGACAGTGTGGCTGACTTCAAGACCTGCTCCAGTCC 6332
 QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAAGGGG 1380
 DB 6333 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAAGGGG 6392
 QY 1381 GTCTGCGGGGAGAGCGCATATGACAGACCACTGCTCATGTGAGACAGATACCGGA 1440
 DB 6393 GTCTGCGGGGAGAGCGCATATGACAGACCACTGCTCATGTGAGACAGATACCGGA 6452
 QY 1441 CATGTCAAAACGCTTCCATGAGATCGTGGGCTTAAAGCTGTATGAATGTGGCAT 1500
 DB 6453 CATGTCAAAACGCTTCCATGAGATCGTGGGCTTAAAGCTGTATGAATGTGGCAT 6512
 QY 1501 GGAACATTCGCGATCAAGCATACACAGAGGCGCTGACGCGCTCCCGAGGCCAAG 1560
 DB 6513 GGAACATTCGCGATCAAGCATACACAGAGGCGCTGACGCGCTCCCGAGGCCAAG 6572
 QY 1561 TATTCAGGCGCTGTGCGGGTGGCTGCTGAGAGTACGTGAGGTTACGCGGGTGGG 1620
 DB 6573 TATTCAGGCGCTGTGCGGGTGGCTGCTGAGAGTACGTGAGGTTACGCGGGTGGG 6632
 QY 1621 GATTTCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCCCTGCGAGTTCCA 1680
 DB 6633 GATTTCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCCCTGCGAGTTCCG 6692
 QY 1681 GCCCGGAAATTTCTCAACGAATGATGGGGGCGGCTGACAGATAGCTCCGGGTCG 1740
 DB 6693 GCCCGGAAATTTCTCAACGAATGATGGGGGCGGCTGACAGATAGCTCCGGGTCG 6752
 QY 1741 AAACCTCTCTACGGAAGAGTCAATTCAGTGGGCTCAACCAATACCTGTTGGG 1800
 DB 6753 AAGCTCTCTACGGAAGAGTCAATTCAGTGGGCTCAACCAATACCTGTTGGG 6812
 QY 1801 TCGAGCTCCATGCGAGCGGCGCAACCGAGTGTAGAGTCTCATCTGCTCACGAC 1860
 DB 6813 TCGAGCTCCATGCGAGCGGCGCAACCGAGTGTAGAGTCTCATCTGCTCACGAC 6872
 QY 1861 CCTTCCCATATCAACGAGAGCGGCTAAGCGGAGCTGCGGAGGCTCTCCCGCTCC 1920
 DB 6873 CCTTCCCATATCAACGAGAGCGGCTAAGCGGAGCTGCGGAGGCTCTCCCGCTCC 6932
 QY 1921 TTGGCAGCTTTCAGCTAGCCAGTGTCTGCGGCTTCTCGAAGGCGCATATTAAC 1980
 DB 6933 TTGGCAGCTTTCAGCTAGCCAGTGTGTCTGCGGCTTCTCGAAGGCGCATATTAAC 6992
 QY 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAATCTCTGTGGCGGATGAGAT 2040
 DB 6993 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAATCTCTGTGGCGGATGAGAT 7052
 QY 2041 GGC 2043
 DB 7053 GGC 7055

RESULT 11
 AF165060 9379 bp RNA linear VRL 04-NOV-1999
 LOCUS Hepatitis C virus strain MD8-2 complete genome.
 DEFINITION
 ACCESSION AF165060
 VERSION AF165060.1 GI:5918958
 KEYWORDS
 SOURCE Hepatitis C virus
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus.
 REFERENCE 1 (bases 1 to 9379)

AUTHORS Nagayama,K., Kurosaki,M., Enomoto,N., Maekawa,S.Y., Miyasaka,Y.,
TITLE Tacawa,J., Izumi,N., Marumo,F. and Sato,C.
JOURNALS Time-related changes in full-length hepatitis C virus sequences and
MEDLINE hepatitis activity
PUBMED Virology 263 (1), 244-253 (1999)
20013325
2 (bases 1 to 9379)
10544098
Nagayama,K., Kurosaki,M., Enomoto,N., Maekawa,S., Miyasaka,Y.,
AUTHORS Sakamoto,N., Fukuma,T., Tacawa,J., Izumi,N., Marumo,F. and Sato,C.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-1999) Second Department of Internal Medicine,
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
Tokyo 113-8519, Japan
Location/Qualifiers
1..9379
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD8-2"
/db_xref="taxon:11103"
/note="from a patient with persistently high serum ALT
values; patient 8, point 2
genotype: 1b"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="AAD56195.1"
/db_xref="GI:5918959"
/translation="MSTNPQRKTKRNTNRPDVRFPGGGQIVGVLLPRRRL
GVRAKTKSERSQPRGRQPIPKARPRGRITAAQVGYWPLYGNMGMAWGLSPG
SRPWGPNDRSRNRNLKVIDITLGTGFDLNGYIPLVAGAGARLAAHGVRLAD
SNVATGM,PGCSFSIPLIALSLCLTIPASAEVNVAGVHVDLNSNSIVYEAAD
IIMHPGCTPCVREGNISRCWALPTLAARNISVPTTIRHVDLVTGTAFCAMV
VGDLCGSEVLVSOLFSPSRHETVQDCNSIYPGHVGRHAAAMMMNSPTALVV
SOLAIRPAVDMVAGAGVAGIAGLAAYSMGNMAKXVLIWMLFPGVLTHTVGAG
AHSATLSPFTVPSQKIQLVNTNGSHIRNTALNCDSLQTLGTLVATFTHRNAG
GCEPMASCRPLDITFDQMGPIYTAEPNSDPDPCYCMHPCGIEPSKCGGVYC
FPPSPVNVGTTDRFGVPTPYMNGENETDVLNLTNRPPGCMWGCPTMNSGTGCTCG
PPCNIGAGNNTLCTPDCFRKHPATYKCGSGMPLPRCLVDYPLMHYPCVNF
TIFKRYMVGVEHRLSACNWTGRCRGLDEDRSELPLSLSTEMOILPSCSTTL
PALSTGLIHRNIVDVOYLVGSGSAVSIARMEVYLLPLIADAVCCIMMML
IAQEAALENVLVINAASVAGHGLISPLVFCAWYIKGRVPGAAYALYVWMLL
LLALPPRAYAMREMAASCGAVFGLAFILTSPYKVFRLIMVQIPIYTRBAH
LQWVPLPNVRGDAIILMCMVHBEILFDITKLILAIFGLMAHQITRVPFVR
AHGILRACMLVAKVAGHYVMAFYKALALGTYYDHLAFLQMAHSGLRDLAVAE
PVFSDMETKIIITGADTAAAGDIIILGLPVASRGRBEILLAGADSLGEMFLAIP
AYSOOTRGLICITISLTGRDKNOYGEVQSVTANOSLALCIVTCMVTVHAGSK
TLAGKRPITQMTYTVDDIVYQWAPRGRSMTPTCTCGSSDLIYVTRHADYPVRRG
DGRGLSPRPVSYLKSGSGGPLCPSGHAAGI FRAVCTRGVAKAVPFIYESMEIT
MNSPVFTDNSPPAVQTFQVAMHAPTSQKSTVPAAYAAQYKVLVNSVAILT
FGAYMSKAYGTDPMIRITGAPITVSTYKFLADGCGSGAADIILICDECHS
TDSITLIGITVLDQAEVAGRLVLAATPPGSAVTPHPNIEEVALSTGSIPIYGK
AIPETVIGGRHLIFCHSKKCDLAATKSLAGLNAVYRGLDVSVIPISGDVVVA
TDALMTGTTGDDSVIICNTCTCTVTVDSLDPTTILETTTVQDVAVSQRGRGRG
REGIYRFTVTPGRPSGMDPSVLCECYDCAWYELTPELTSIRAKAYITGLVQC
DHLBEPWSEFTGLTHIDAFLSQTKOAGNPFYVAYQVTCARQAPPSDDWKC
LIRLKPITLHGPPLIRLGAQNEVTLPIITKIFIMACASDLAEVTVGTVLVGLVA
ALAAVCLTSSVIVIRIILSGPAVILPREVLYOBEPMESCAEHLPIBEGMOLAB
OPKOKALGILQTAQKQAEAAAVFVSEKRALETPMAKMMNITISQIYIAGISLTPGN
PAISLMFTASTITSPILTQTTIMRNILGAWAQLAPSAASAVIAGIAGAAVGI
GAGKLVLDILAGVAGVALAFKVMGSEMTEDLILPAILISPGALVGVCA
LILRVHGEQEGAVQMMNRIAPASGNHVSPTHYVESPAARVQIISLITVGLK
RLHWINEDCTSPGSGMWLDVMDICTVLDFTKMLQSGIKMPRLPGFISLCOQYK
AGWRGDIMQTCPCGAOITGHVKNKSNIVGPKTCSNWHGCTPPLNATVTCPSRP
APVSRALMRVAAEEVYVTVRGDQPHVYKMTQNTVNCQCPAPBEPFENDVGLHR
YAPACKPLIREVITFOVGLNQYLVGSQLPCEPPEVAVITSLMTPSHITATARL
AAGSPSLASSASQSLASPLAKATCTTHDSFDADLIENTLMRMEGNTIRVSEN
KVIILISFDPLRAEBDESEVSAALIRSRKFPALPWAPNDPNPLBEMKPDY
VPPVHAGCPLEPTKAPIPPRKRITVLTSSVSALEIATKTFGSGSADVSGT
ATAPDQASDDGKSDVESYSGMPLBEGPDPLSDSMSTVEBBSSEDVVCMS
YTMTCALITPCAAESKILPINLSLHLHHMYVATTSRSASLQOKYTPRLQYLD
DHYRVLAKMKKASTVAKLUSIEACKLTPHSARSFEGADGVSLSKAANH
RSWBDLLDTEPTIDTITMAKSEVFCVQPEKGRKPARLIYFPLGAVRCEKMAIYD

FEATURES
source
CDS

ORIGIN		Query Match	88.0%; Score 1798.2; DB 14; Length 9379;
Beet Local Similarity		92.5%; Pred. No. 0;	
Matches 1890; Conservative		0; Mismatches 153; Indels 0; Gaps 0;	
QY	1	TGGAGGGCGTCTTCACAGGCTTACCCACGCTGATGCGCACTTCTGCCAACAAG	60
DB	5001	TGGAGAGTGTCTTACAGGCTTACCCACATAGATGCCACTTCTGCCAACAAG	5060
QY	61	CAGGACAGACAACTTCCCTTACCTGGTGGCGTACAGGCTACTGTGGCTTAGGGCC	120
DB	5061	CAGGACAGAGTAACTTCCCTTACCTGGTGGCGTACAGGCTACTGTGGCGCGGCT	5120
QY	121	CAGGCCACCTCCATCATGAGGATCAATGTGGAAGTCTCATAGGCTAAAGCTACT	180
DB	5121	CAGGCTCCACCTCCATCTGGGACCAATGTGGAAGTCTCATAGGCTAAACCTACA	5180
QY	181	CTGCGCGGGCCAAACCCCTTGTGTATAGGCTGGAGCCGTCCAAACGAGTCACTC	240
DB	5181	CTGCACGGGCCAAACCCCTGTGTATAGGCTAGAGCCGTCCAAATGAGGTCACTCC	5240
QY	241	ACACACCCCATPACCAATTATCATGTGCATATGTCAGCCGACCTGGAAGTCCTCACG	300
DB	5241	ACACACCCCATPACCAATTATCATATGCAATGCAATGCGTGAACCTGGAAGTCCTCAC	5300
QY	301	AGCACTGGAGTGTGTGGCGGGGTCTTGGAGCTTGGCTGCTGCTTGTGCAACA	360
DB	5301	AGCACTGGAGTGTGTGGCGGGGTCTTGGAGCTTGGCTGCTGCTGCTGCAACA	5360
QY	361	GCGACGCTGATCTTGTGGTGAAGATCATCTTGTCCGGCGCGGCTATTTGTTCCGAC	420
DB	5361	GCGACGCTGATCTTGTGGCGAGATTCATCTTGTCCGGGAAGCGGCTGTCATCCGAC	5420
QY	421	AGGAAAGTCTCTACAGAGATTCATGATGGAAGTGGCGCTCCACCTCCCTTAC	480
DB	5421	AGGAAAGTCTCTACAGAGATTCATGATGGAAGTGGCGCTCCACCTCCCTTAC	5480
QY	481	ATCGACAGGGAATCAGCTCGCCGAGCATTCAGCAAAAAGCGCTCGGTTGTGACG	540
DB	5481	ATCGAACAGGAATCAGCTCGCCGAGCATTCAGCAAAAAGCGCTCGGTTGTGACA	5540
QY	541	ACAGCCACCAAGCAAGCGAGCGCTCTCCCGTGTGTGAGTCCAAGTGGCAGCCCTT	600
DB	5541	ACAGCCACCAAGCAAGCGAGCGCTCTCCCGTGTGTGAGTCCAAGTGGGAGCCCTC	5600
QY	601	GAGACCTTCTGGGCGAACAATGTGGAATTCATACCGGGATACAGTATTACAGGC	660
DB	5601	GAGACCTTCTGGGCGAACAATGTGGAATTCATACCGGGATACAGTATTACAGGC	5660
QY	661	TTGTCCATCTGCGGGGAATCCGCGATGTCATGATGAGTGGCTTACAGACTCTGTG	720
DB	5661	TTGTCCATCTGCGGGGAATCCGCGATGTCATGATGAGTGGCTTACAGACTCTCAT	5720
QY	721	ACTAGCCCGCTACCAACCAATCTACCTCTGCTTAAATCTCTGGGGGAGTGGTAGCC	780
DB	5721	ACTAGCCCGCTACCAACCAATCTACCTCTGCTTAAATCTCTGGGGGAGTGGTAGCT	5780
QY	781	GCCCACTGGCTCCCCCAAGTCTCTTCAAGTGGCGCGCATTTCTGTGTG	840
DB	5781	GCCCACTGGCTCCCCCAAGTCTCTTCAAGTGGCGCGCATTTCTGTGTG	5840
QY	841	GCTGTGGCAGATGAGCTTGGGAAGTGTGTGACATCTTGGCGGGCTATGAGACA	900
DB	5841	GCTGTGGCAGATGAGCTTGGGAAGTGTGTGACATCTTGGCGGGCTATGAGGACA	5900

QY 901 GGAGTGGCAGGCGCGCTCGTGGCCCTTTAAGTCATAGAGCGCGGAAATGCCCTCCACCGAG 960
 DB 5901 GGGGGTGGCTGGGCGCACTCGTGGCTTTTAAGTCATAGAGCGCGGAAATGCCCTCCACCGAG 5960
 QY 961 GACCTGTGTTAACTTACTCCCTGCGCATCTCTCTCTCTGTTGCTCGTGGGATCGTG 1020
 DB 5961 GACCTGTGTTAACTTACTCCCTGCGCATCTCTCTCTCTGTTGCTCGTGGGATCGTG 6020
 QY 1021 TGGCGAGGATACCTGCTGCGGACGTTGGGTCCAGGGGAGGGGCTGTGAGTGAAGAAC 1080
 DB 6021 TGGCGAGGATACCTGCTGCGGACGTTGGGTCCAGGGGAGGGGCTGTGAGTGAAGAAC 6080
 QY 1081 CGGCTGATAGCTTGGCTCGCGGGGTAACTGTTTCCCGGACGACTATGTCAGAG 1140
 DB 6081 CGGCTGATAGCTTGGCTCGCGGGGTAACTGTTTCCCGGACGACTATGTCAGAG 6140
 QY 1141 AGCGACGCGCGAGCAGTGTCACTCAGATCTCTCCGACCTTAACTATCACTCACTGTTG 1200
 DB 6141 AGCGACGCGCGAGCAGTGTCACTCAGATCTCTCCGACCTTAACTATCACTCACTGTTG 6200
 QY 1201 AAGAGGCTCCACCACTGATTTAAACGAGAGCTGCTCCAGCCCTGCTCGGCTGTGCTA 1260
 DB 6201 AAGAGGCTCCACCACTGATTTAAACGAGAGCTGCTCCAGCCCTGCTCGGCTGTGCTA 6260
 QY 1261 AAGGATGTTTGGGACTGATATGCAAGTGTGCTGACTTGAAGAAGCTGGCTCCAGTCC 1320
 DB 6261 AAGGATGTTTGGGACTGATATGCAAGTGTGCTGACTTGAAGAAGCTGGCTCCAGTCC 6320
 QY 1321 AAGCTCTGCTCGCGCATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAAGGGG 1380
 DB 6321 AAGCTCTGCTCGCGCATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAAGGGG 6380
 QY 1381 GTCTGGCGGGGAGAGCGGCTCATGCAAGACCACTGCTCATGTGAGACAGATCAACCGA 1440
 DB 6381 GTCTGGCGGGGAGAGCGGCTCATGCAAGACCACTGCTCATGTGAGACAGATCAACCGA 6440
 QY 1441 CATGTCAAAAAAGTTCATGAGGATCGTGGGCTTAAAGCTGTATGAACATGTTGGCAT 1500
 DB 6441 CATGTCAAAAAAGTTCATGAGGATCGTGGGCTTAAAGCTGTATGAACATGTTGGCAT 6500
 QY 1501 GGAACATTTCCCATCAACGATACCAACGAGGCGCCCTGACGCGCTCCCGAGCGCCAAAC 1560
 DB 6501 GGAACATTTCCCATCAACGATACCAACGAGGCGCCCTGACGCGCTCCCGAGCGCCAAAC 6560
 QY 1561 TATTCAGGCGCGCTGTGGCGGTGCTGCTGAGAGATACGTGAGGTTACGCGGCTGGGG 1620
 DB 6561 TATTCAGGCGCGCTGTGGCGGTGCTGCTGAGAGATACGTGAGGTTACGCGGCTGGGG 6620
 QY 1621 GATTTCACTAGTACGAGGATGACCACTGACCAAGTAAATGCCCGTGGCAGAGTTCA 1680
 DB 6621 GATTTCACTAGTACGAGGATGACCACTGACCAAGTAAATGCCCGTGGCAGAGTTCA 6680
 QY 1681 GCGCCCGAATTTCTTACAGAAATGATGAGGTGCGGCTGACAGGATACGCTCGGAGTGC 1740
 DB 6681 GCGCCCGAATTTCTTACAGAAATGATGAGGTGCGGCTGACAGGATACGCTCGGAGTGC 6740
 QY 1741 AAACCTCTTCTTACGAGGAGGATCAATTCAGGCTGGGCTCAACCAATACCTGTGGG 1800
 DB 6741 AAACCTCTTCTTACGAGGAGGATCAATTCAGGCTGGGCTCAACCAATACCTGTGGG 6800
 QY 1801 TCGGAGCTCCCATGAGGCGGACCGAGTGAAGAGTGTGCACTTCCATGCTCACCGAC 1860
 DB 6801 TCGGAGCTCCCATGAGGCGGACCGAGTGAAGAGTGTGCACTTCCATGCTCACCGAC 6860
 QY 1861 CCTTCCTCATCAGCAGAGAGCGCTAAGCGAGGCTGACAGGAGGCTCTCCCTCCCTCC 1920
 DB 6861 CCTTCCTCATCAGCAGAGAGCGCTAAGCGAGGCTGACAGGAGGCTCTCCCTCCCTCC 6920
 QY 1921 TTGGCAGGCTTTCAGCTAAGCAAGTGTGCGCTTCTCGAAGGCGCATATCATTACC 1980
 DB 6921 TTGGCAGGCTTTCAGCTAAGCAAGTGTGCGCTTCTCGAAGGCGCATATCATTACC 6980

QY 1981 CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGATGATG 2040
 DB 6981 CATCATGACTCCCGGAGGCTGACCTCATGAGGCCAACCTCTGTGGCGGATGATG 7040
 QY 2041 GGC 2043
 DB 7041 GGC 7043
 RESULT 12
 AR005087
 LOCUS AR005087 7863 bp DNA linear PAT 04-DEC-1998
 DEFINITION Sequence 35 from patent US 5747339.
 ACCESSION AR005087
 VERSION AR005087.1 GI:3965966
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1 (bases 1 to 7863)
 AUTHORS Okayama,H., Fuke,I., Mori,C., Takamizawa,A. and Yoshida,I.
 TITLE Non-A, non-B hepatitis virus genomic CDNA and antigen polypeptide
 JOURNAL Patent: US 5747339-A 35 05-MAY-1998;
 FEATURES
 source
 1..7863
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Query Match 87.9%; Score 1796.6; DB 6; Length 7863;
 Best Local Similarity 92.5%; Pred. No. 0;
 Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
 QY 1 TGGAGAGGCGCTTTTCAACGAGGCTCAACCACTGATGCCACTTCTGCTCCAAACAAAG 60
 DB 3505 TGGAGAGGCGCTTTTCAACGAGGCTCAACCACTGATGCCACTTCTGCTCCAAACAAAG 3564
 QY 61 CAGGAGAGGAGCAACTTCCCTTACCTGGTGGGCTGACGAGCTACTGTGCGCTAGGGCC 120
 DB 3565 CAGGAGAGGAGCAACTTCCCTTACCTGGTGGGCTGACGAGCTACTGTGCGCTAGGGCC 3624
 QY 121 CAGGCCCACTTCATCATGATGATCAATGTGGAAGTGTCTCATACGCTAAAGCTTACT 180
 DB 3625 CAGGCCCACTTCATCATGATGATCAATGTGGAAGTGTCTCATACGCTAAAGCTTACT 3684
 QY 181 CTGCGGCGGCCCAACACCTTGTCTGTATAGGCTGGAGCCCTGCAAAACGAGTCACTTC 240
 DB 3685 CTGCGGCGGCCCAACACCTTGTCTGTATAGGCTGGAGCCCTGCAAAACGAGTCACTTC 3744
 QY 241 ACACACCCCATTAACCAATTCATCATGAGGATGATGACGAGCCTGAGGAGTGGTCAAG 300
 DB 3745 ACACACCCCATTAACCAATTCATCATGAGGATGATGACGAGCCTGAGGAGTGGTCAAG 3804
 QY 301 AGCACCCTGGTCTGATGAGGCGGAGTCTTGCAGCTGTGAGTGTGATTTGTTGACACA 360
 DB 3805 AGCACCCTGGTCTGATGAGGCGGAGTCTTGCAGCTGTGAGTGTGATTTGTTGACACA 3864
 QY 361 GGCAGGCTGTATTGTGGTATGATCATCTTGTCCGAGCGGCTGATTTGTTCCGAC 420
 DB 3865 GGCAGGCTGTATTGTGGTATGATCATCTTGTCCGAGCGGCTGATTTGTTCCGAC 3924
 QY 421 AGGGAAGTCTTACACAGAGTTCATGATGAGAAAGTGTGGCGATCGACCTCCCTTAC 480
 DB 3925 AGGGAAGTCTTACACAGAGTTCATGATGAGAAAGTGTGGCGATCGACCTCCCTTAC 3984
 QY 481 ATCGAGCAGGAGTACAGCTCGCGAGCAGTTCAAGCAAAAAGCGCTCGGTTGTGTCAG 540
 DB 3985 ATCGAGCAGGAGTACAGCTCGCGAGCAGTTCAAGCAAAAAGCGCTCGGTTGTGTCAG 4044
 QY 541 ACAGCACCACCAAGCGAGGCGCTGCTCCGTTGTGTGAGTCCAATGTGCGAGCCTTT 600
 DB 4045 ACAGCACCACCAAGCGAGGCGCTGCTCCGTTGTGTGAGTCCAATGTGCGAGCCTTT 4104

```

OY 601 GAGACCTTGTGGGCGAACAATGTGAACTTCAATCAGCGGATACATGACTTACAGGC 660
Db 4105 GAGACATTTCTGGGCGAAGACATGTGAAATTTCAATCAGCGGATACAGTACTTACAGGC 4164
OY 661 TTGTCCACTGTGCGGGGAATCCGGGATGTGCATGATGAGCGCTTCAACAGCTCTGTC 720
Db 4165 TTATCACTTGTCTGGGAACCCCGCATAGCATTTGATGGCATTCACAGCTCTATC 4224
OY 721 ACTAGCCGCTCACCACCCAAATCTACCTCTGTCTTAAATCTGCGGGGATGGGTAGCC 780
Db 4225 ACCAGCCGCTCACCACCCAAAGTAACCTCTGTCTTAAATCTTGGGGGGGTGGGTGGCT 4284
OY 781 GCCCACTGTCTCCCCCAAGTCTGTGATTTTGTGAGGGCGCGCATCTCTGTGTCG 840
Db 4285 GCCCACTGTCTCCCCCAAGTCTGTGATTTTGTGAGGGCGCGCATCTCTGTGTCG 4344
OY 841 GCTGTGGAGAGATAGGCTTGGGAAGGTGTGTGACATCTTGGCGGGATGTGAGCA 900
Db 4345 GCTGTGGAGAGATAGGCTTGGGAAGGTGTGTGACATCTTGGCGGGATGTGAGCA 4404
OY 901 GGAGTGGAGAGCGCGCTGTGACCTTTAAGTTCATAGAGCGGAAATGCCCTTCCACGAG 960
Db 4405 GGAGTGGAGAGCGCGCTGTGACCTTTAAGTTCATAGAGCGGAGATGCCCTTCCACGAG 4464
OY 961 GACCTGTGTTAATTAATCTCTGCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 4465 GACCTGTGTTAATTAATCTCTGCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 4524
OY 1021 TGGCGAGGATATCTGCTGCGGACGATGGGTCCAGGGGAGGGGGTGTGAGTGAAGCA 1080
Db 4525 TGTGAGCAATATCTGCTGCGGACGATGGGTCCAGGGGAGGGGGTGTGAGTGAAGCA 4584
OY 1081 CGGCTGATAGCTTGTGCGCTGTGCGGGGTAAACATGTTTCCCGCACGACATGTGTCAGAG 1140
Db 4585 CGGCTGATAGCTTGTGCGCTGTGCGGGGTAAATGTTTCCCGCACGACATGTGTCAGAG 4644
OY 1141 AGCGAGCGCGGACGACGTGTCTACTAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 4645 AGCGAGCGCGGACGACGTGTGTACTAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTG 4704
OY 1201 AAGAGGCTCCACCAATGATTAACGAGATGTGCTCCAGCGCTGCGGGCTGTGAGCTGA 1260
Db 4705 AAGAGGCTCCACCAATGATTAACGAGATGTGCTCCAGCGCTGCGGGCTGTGAGCTGA 4764
OY 1261 AGGAGTGTGGAGATGATATGACAAGTGTGCTGATTAAGACCTGCTCCAGTCC 1320
Db 4765 AGGAGTGTGGAGATGATATGACAAGTGTGCTGATTAAGACCTGCTCCAGTCC 4824
OY 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGTCCAAAGTGGTAAACAGGGG 1380
Db 4825 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGTCCAAAGTGGTAAACAGGGG 4884
OY 1381 GTCTGCGGGGAGACGCGATCATGAGACACCTGTCTATGTGAGCAAGATCACCGGA 1440
Db 4885 GTCTGCGGGGAGACGCGATCATGAGACACCTGTCTATGTGAGCAAGATCACCGGA 4944
OY 1441 CATGTCAAAAACGTTCCATGAGATCGTTGGGCTTAAGACCTGTATGAATGTGGCAT 1500
Db 4945 CATGTCAAAAACGTTCCATGAGATCGTTGGGCTTAAGACCTGTATGAATGTGGCAT 5004
OY 1501 GGAACATTTCCCATCAACGATACACACGCGCCCTGTGACGCGCTCCCGAGGCCAAG 1560
Db 5005 GGAACATTTCCCATCAACGATACACACGCGCCCTGTGACGCGCTCCCGAGGCCAAG 5064
OY 1561 TATTCAGGGGCTGTGAGGGGTGTGCTGAGAGATGAGTGAAGTTTACGCGGGGTGGG 1620
Db 5065 TATTCAGGGGCTGTGAGGGGTGTGCTGAGAGATGAGTGAAGTTTACGCGGGGTGGG 5124
OY 1621 GATTTCACTAAGTGAAGCATGACCACTGACCAAGTAAATGGCCGTGACAGTTTCCA 1680
Db 5125 GATTTCACTAAGTGAAGCATGACCACTGACCAAGTAAATGGCCGTGACAGTTTCCG 5184
OY 1681 GCCCCGAATTTCTCAAGAGTGAAGTGGGGTGGGCTGACACAGTACGCTCCGCGCTGC 1740

```

```

Db 5185 GCTCTGAATTTCTTCCGAGGTGAGCGAGTGCCTTGCACAGGTACGCTCCGCGCTGC 5244
OY 1741 AAACCTCTCTTACGGGAGAGAGTGCATTTCCAGTGTGGCTCAACCAATATCTGTGGG 1800
Db 5245 AGGCTCTCTTACGGGAGAGAGTGCATTTCCAGTGTGGCTCAACCAATATCTGTGGG 5304
OY 1801 TCGAGCTTCCATGCGAGCCCGAACCAGATGTAGAGTGTCTCATCTTCCATGTCACGAC 1860
Db 5305 TCAGAGTACATGAGACCCGAAACCGGATGTAGAGTGTCTCATCTTCCATGTCACGAC 5364
OY 1861 CCTCCCATCAAGACGAGACGCGCTAAGGAGCGCTGAGCGAGGGGTCTCCCTCC 1920
Db 5365 CCTCCCATCAAGACGAGAAACGCTAAGGAGTGTGAGCGAGGGGTCTCCCTCC 5424
OY 1921 TTGGCAGCTCTTACGATGACGATGTGTGTGCTGCTGCTGCTGAGAGCGACATATTAAC 1980
Db 5425 TTGGCAGCTCTTACGATGACGATGTGTGTGCTGCTGCTGCTGAGAGCGACATATTAAC 5484
OY 1981 CAAATGACTTCCAGACGCTGACCTCATTCAGAGCCAACTTCTGTGGCGCATGAGATG 2040
Db 5485 CAAATGACTTCTCCGACGCTGACCTCATTCAGAGCCAACTTCTGTGGCGCATGAGATG 5544
OY 2041 GGC 2043
Db 5545 GGC 5547

RESULT 13
AR064526 7863 bp DNA linear PAT 29-SEP-1999
LOCUS AR064526
DEFINITION Sequence 35 from patent US 5847101.
ACCESSION AR064526
VERSION AR064526.1 GI:5993834
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 7863)
AUTHORS Okayama,H., Fuke,I., Mori,C., Takamizawa,A. and Yoshida,I.
TITLE Non-A, non-B hepatitis virus genomic cDNA and antigen polypeptide
JOURNAL Patent: US 5847101-A 35 08-DEC-1998;
FEATURES
source
1..7863
/organism="Unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 87.9%; Score 1796.6; DB 6; Length 7863;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

OY 1 TGGGAGGGGCGTCTTACAGAGGCTTCAACCGACGATGAGCCACTTCTGTCCAAACAAAG 60
Db 3505 TGGGAGAGTGTCTTACAGGCTTCAACCAATATGATGACACTTCTGTCCAGACCAAG 3564
OY 61 CAGGACGAGACAACTTCCCTTACCTGTGTGGGTACAGGCTACTGTGTGCGCTAAGGCC 120
Db 3565 CAGGACGAGACAACTTCCCTTACCTGTGTGAGTATCAACAGCAACGATGTGCGCGGCT 3624
OY 121 CAGGCCCCACCTCCATCATGAGATCAAAATGTGGAAGTGTCTTCAATCGCTAAAGCTACT 180
Db 3625 CAGGCCCCACCTCCATCATGAGATCAAAATGTGGAAGTGTCTTCAATCGCTAAAGCTACT 3684
OY 181 CTGGCGGGGCGCAACACCTTGTGTATATGAGTGTGGAGCGGTTCACAAAGAGGTACCTTC 240
Db 3685 CTGGCGGGGCGCAACACCTTGTGTATATGAGTGTGGAGCGGTTCACAAAGAGGTACCTTC 3744
OY 241 ACACACCCCATTAACCAATTCATCATGTGACATGATGACCGGACCTGAGAGGTGTGACG 300
Db 3745 ACCACCCCATTAACCAATTCATCATGTGACATGATGACCGGACCTGAGAGGTGTGACG 3804
OY 301 AGCACTGGGTGCTGTGTGGCGGGGTCTTGGACGCTTGAGCTTGATTTGCTTGACACACA 360

```


Db 3505 TGGGAGAGTGTCTTACAGGGCTCACCCATATAGACACTTCTTGTCCAGACCAAG 3564
Qy 61 CAGGAGGAGACAACTCTCCCTACCTGTGTGCGTACAGGCTATGTGTGCGTAAAGGC 120
Db 3565 CAGGAGGAGACAACTCTCCCTACCTGTGTGCGTACAGGCTATGTGTGCGTAAAGGC 3624
Qy 121 CAGGAGGAGACAACTCTCCCTACCTGTGTGCGTACAGGCTATGTGTGCGTAAAGGC 180
Db 3625 CAGGAGGAGACAACTCTCCCTACCTGTGTGCGTACAGGCTATGTGTGCGTAAAGGC 3684
Qy 181 CTGGCGGGGGCCAAACCTTGTGTATAGGCTGGGAGCGCTCCAAACGAGGTCAACCTC 240
Db 3685 CTGGCGGGGGCCAAACCTTGTGTATAGGCTGGGAGCGCTCCAAACGAGGTCAACCTC 3744
Qy 241 ACACACCCCATTAACAAATTCATCATGTGCATGTCAAGCCGACCTGGAGGTCTGACG 300
Db 3745 ACCACCCCATTAACAAATTCATCATGTGCATGTCAAGCCGACCTGGAGGTCTGACG 3804
Qy 301 AGCACCTGGGTGTGTGTGGGCGGGGTCTTGGCAGCTTGGCTGCTATTTGCTTGAACA 360
Db 3805 AGCACCTGGGTGTGTGTGGGCGGGGTCTTGGCAGCTTGGCTGCTATTTGCTTGAACA 3864
Qy 361 GGCACGCTGTATGT 420
Db 3865 GGCACGCTGTATGT 3924
Qy 421 AGGGAAGTCTCTTACAGGAGTTCATGATGATGAAAGTGTGCGTGTGCACTCCCTTAC 480
Db 3925 AGGGAAGTCTCTTACAGGAGTTCATGATGATGAAAGTGTGCGTGTGCACTCCCTTAC 3984
Qy 481 ATCGACAGGGAATGAGTGTGCGCCGACGATTCAAGCAAAAGCGCTGGGTGTGTCGAG 540
Db 3985 ATCGACAGGGAATGAGTGTGCGCCGACGATTCAAGCAAAAGCGCTGGGTGTGTCGAG 4044
Qy 541 ACAGCACCAAGCAAGCGGAGCGCGCTGTCCGCTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db 4045 ACAGCACCAAGCAAGCGGAGCGCGCTGTCCGCTGTGTGTGTGTGTGTGTGTGTGTGT 4104
Qy 601 GAGACCTTCTGGGGCAAAACATGTGTGAATTCATCAGCGGATACAGTACTTACAGGC 660
Db 4105 GAGACCTTCTGGGGCAAAACATGTGTGAATTCATCAGCGGATACAGTACTTACAGGC 4164
Qy 661 TTGTCACTCTGCTGTGGGAATCCCGGATGTGATCACTGATGCGGTTCACAGCTCTGTC 720
Db 4165 TTGTCACTCTGCTGTGGGAATCCCGGATGTGATCACTGATGCGGTTCACAGCTCTGTC 4224
Qy 721 ACTAGCCGCTACACCCCAATCTACCCCTGCTTAACTCCTGGGGGATGGGTAGGC 780
Db 4225 ACTAGCCGCTACACCCCAATCTACCCCTGCTTAACTCCTGGGGGATGGGTAGGC 4284
Qy 781 GCCCACTGCTCCCGCAGTGTGCTTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 840
Db 4285 GCCCACTGCTCCCGCAGTGTGCTTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 4344
Qy 841 GCTGTGTGAGCAATAGGCTTGGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 900
Db 4345 GCTGTGTGAGCAATAGGCTTGGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 4404
Qy 901 GGAAGTGGCAGGCGCGCTG 960
Db 4405 GGAAGTGGCAGGCGCGCTG 4464
Qy 961 GACCTGTAACTTACTCTGCTGCACTCTCTCTGCTGTGCTGTGTGTGTGTGTGTGTGTG 1020
Db 4465 GACCTGTAACTTACTCTGCTGCACTCTCTCTGCTGTGCTGTGTGTGTGTGTGTGTGTG 4524
Qy 1021 TGGCAGCGATATCTG 1080
Db 4525 TGGCAGCGATATCTG 4584
Qy 1081 CGGCTGATAGCTG 1140
Db 4585 CGGCTGATAGCTG 4644

Qy 1141 AGCGAGCGCGGAGCAGTGTCACTGATTCCTCTTCCGACTTATCATCAACCACTGTG 1200
Db 4645 AGCGAGCGCGGAGCAGTGTCACTGATTCCTCTTCCGACTTATCATCACTGATTCG 4704
Qy 1201 AAGAGCTCCCAACAGTGTATTAACGAGGACTGTCCACGCCCTGTCCGGCTGTGCTA 1260
Db 4705 AAGAGCTCCCAACAGTGTATTAACGAGGACTGTCCACGCCCTGTCCGGCTGTGCTA 4764
Qy 1261 AAGAGTGTGTGGAGCTGTGATATGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1320
Db 4765 AAGAGTGTGTGGAGCTGTGATATGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 4824
Qy 1321 AAGCTCTGCGCGATTACCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1380
Db 4825 AAGCTCTGCGCGATTACCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 4884
Qy 1381 GTTGTGGGGGGAGCAGGATCATGACAGCAACCTGTGATGTGAGCAGATCAACCGGA 1440
Db 4885 GTTGTGGGGGGAGCAGGATCATGACAGCAACCTGTGATGTGAGCAGATCAACCGGA 4944
Qy 1441 CATGTCAAAAACGTTTCCATGAGATCTGTGGGCTTAAGACCTGTATGATGTGAGAT 1500
Db 4945 CATGTCAAAAACGTTTCCATGAGATCTGTGGGCTTAAGACCTGTATGATGTGAGAT 5004
Qy 1501 GGAACATTCCTCCATCAACGATCAACAGCGGCTCTGACAGCTTCCCAAGCGCAAC 1560
Db 5005 GGAACATTCCTCCATCAACGATCAACAGCGGCTCTGACAGCTTCCCAAGCGCAAC 5064
Qy 1561 TATTTCAAGGGGCTGTGTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1620
Db 5065 TATTTCAAGGGGCTGTGTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 5124
Qy 1621 GATTTCACTACGTGTACAGATGATGACACTGACCAAGTAAATGCGCTGTGCAAGTTCA 1680
Db 5125 GATTTCACTACGTGTACAGATGATGACACTGACCAAGTAAATGCGCTGTGCAAGTTCA 5184
Qy 1681 GCCCGCAATTTCTTCAAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1740
Db 5185 GCCCGCAATTTCTTCAAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 5244
Qy 1741 AAACCTCTCTACGAGGAGAGGTGTATTCAGAGTGTGTGTGTGTGTGTGTGTGTGTG 1800
Db 5245 AAACCTCTCTACGAGGAGAGGTGTATTCAGAGTGTGTGTGTGTGTGTGTGTGTGTG 5304
Qy 1801 TCGCAGCTCCCATGTGCAAGCCCGAACCGGATGTGAGTGTGTGTGTGTGTGTGTGTG 1860
Db 5305 TCGCAGCTCCCATGTGCAAGCCCGAACCGGATGTGAGTGTGTGTGTGTGTGTGTGTG 5364
Qy 1861 CCTTCCCATCAACGAGAGAGCGGCTTAAGCGCAGGCTGTGCAAGGGGTCTTCCCTTCC 1920
Db 5365 CCTTCCCATCAACGAGAGAGCGGCTTAAGCGCAGGCTGTGCAAGGGGTCTTCCCTTCC 5424
Qy 1921 TTGGCAGCTCTTTCAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1980
Db 5425 TTGGCAGCTCTTTCAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 5484
Qy 1981 CAATATGACTTCCAGAGCTGTGACTCATGTGAGGCAACCTTCTGTGTGTGTGTGTG 2040
Db 5485 CAATATGACTTCCAGAGCTGTGACTCATGTGAGGCAACCTTCTGTGTGTGTGTGTG 5544
Qy 2041 GGC 2043
Db 5545 GGC 5547

RESULT 15
ARI45568 7863 bp DNA linear PAT 08-AUG-2001
LOCUS ARI45568
DEFINITION Sequence 35 from patent US 6217872.
ACCESSION ARI45568
VERSION ARI45568.1 GI:15108757
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7863)
AUTHORS Okayama,H., Fuke,I., Mori,C., Takamizawa,A. and Yoshida,I.
TITLE Non-A, non-B hepatitis virus genomic cDNA and antigen polypeptide
JOURNAL Patent: US 6217872-A 35 17-APR-2001;
FEATURES Location/Qualifiers
source 1..7863
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 87.9%; Score 1796.6; DB 6; Length 7863;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 1 TGGAGGCGCTTTCACAGGCTTCAACCCAGTGGATGCCACTTCTGTCCGAAACAAAG 60
DB 3505 TGGAGAGTGTCTTCAAGGCTTCAACCCATATAGATGACACTTGTCTCCAGACAAAG 3564
QY 61 CAGGACGAGACAACTTCCCTTACTGTGGCGTACAGGCTTACTGTGTGCGTGAAGGCC 120
DB 3565 CAGGACGAGACAACTTCCCTTACTGTGGCATACCAAGCCAGTGTGCGCAGAGGCT 3624
QY 121 CAGGCGCCACCTTCATATGGGATCAAAATGTGAAAGTGTCTATACGGCTTAAAGCTTACT 180
DB 3625 CAGGCGCCACCTTCATATGGGATCAAAATGTGAAAGTGTCTATACGGCTTAAAGCTTACT 3684
QY 181 CTGCGCGGCGCAACACCTTGTGTATAGCTGGAGCCGCTCCAAACAGAGTCAACCTTC 240
DB 3685 CTGCGAGGCGCAACACCTTGTGTATAGCTGGAGCCGCTCCAGATAGAGTCAACCTTC 3744
QY 241 ACAACACCCATTAACAATTCATATGATGATGATGATGATGATGATGATGATGATGATG 300
DB 3745 ACCCAACCCATTAACAATTCATATGATGATGATGATGATGATGATGATGATGATGATG 3804
QY 301 AGCAGCTGGTGTCTGTGGCGGCGGCTTTCAGAGCTGTGGCTGTGGCTGTGGCTGTGGCT 360
DB 3805 AGCAGCTGGTGTCTGTGGCGGCGGCTTTCAGAGCTGTGGCTGTGGCTGTGGCTGTGGCT 3864
QY 361 GGCAGCGTGTATGTGGGTAGATCATCTTGTCCGCGCGCGCGCTATTGTTCCGAC 420
DB 3865 GGCAGTGTGTATGTGGGTAGATCATCTTGTCCGCGCGCGCGCTATTGTTCCGAC 3924
QY 421 AGGGAATCTCTTACAGAGATTCATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 3925 AGGGAATCTCTTACAGAGATTCATGATGATGATGATGATGATGATGATGATGATGATG 3984
QY 481 ATCGAGCAGGGAATGCAAGTCCGAGCAGTTCAGACAAAGCGCTCGGTTGCTGACG 540
DB 3985 ATCGAGCAGGGAATGCAAGTCCGAGCAGTTCAGACAAAGCGCTCGGTTGCTGACG 4044
QY 541 ACAGCCACCAACAGCGAGGCGCGTGTCTCCGCTGTGGAGTCCAAAGTGGCGAGCGCTT 600
DB 4045 ACAGCCACCAACAGCGAGGCGCGTGTCTCCGCTGTGGAGTCCAAAGTGGCGAGCGCTT 4104
QY 601 GAGACCTTCTGGGCGAAACACATGTGGAATTCATCAGCGGGAATACATTAAGCAGGC 660
DB 4105 GAGACCTTCTGGGCGAAACACATGTGGAATTCATCAGCGGGAATACATTAAGCAGGC 4164
QY 661 TTGTCACTCTGCGTGGGAATCCCGGATTTGATGATGATGATGATGATGATGATGATGATG 720
DB 4165 TTATCACTCTGCGTGGGAATCCCGGATTTGATGATGATGATGATGATGATGATGATGATG 4224
QY 721 ACTAGCCGCTCACCACCAATCTACCTCTGTCTTAATCATCTGTGGGGAATGGATAGCC 780
DB 4225 ACCAGCCGCTCACCACCAATCTACCTCTGTCTTAATCATCTGTGGGGAATGGATAGCC 4284
QY 781 GCCCACTCTGCTCCCGCGAGTGTCTGAGCTTTTGTGATGAGCGCGCGCATTTGCTGTGG 840
DB 4285 GCCCACTCTGCTCCCGCGAGTGTCTGAGCTTTTGTGATGAGCGCGCGCATTTGCTGTGG 4344

QY 841 GCTGTGTGACAGATAGGCTTGGGAAGTGTGTGTGACATCTTGGCGGCTATGAGCA 900
DB 4345 GCTGTGTGACAGATAGGCTTGGGAAGTGTGTGTGACATCTTGGCGGCTATGAGCA 4404
QY 901 GGAATGGCAGCGCGCTGTGTCTTTAAGTATAGCGCGGAATGCCCTTCACCGAG 960
DB 4405 GGAATGGCAGCGCGCTGTGTCTTTAAGTATAGCGCGGAGATGCCCTTCACCGAG 4464
QY 961 GACCTGTTAATCTTCTCCCTGCAATCCCTCTCTCCGCTGTGCTGTGCTGTGCTGTG 1020
DB 4465 GACCTGTTAATCTTCTCCCTGCAATCCCTCTCTCCGCTGTGCTGTGCTGTGCTGTG 4524
QY 1021 TGGCAGCAGATACTGCTGTGCGACGCTGGTCCAGGGAAGGAGGCTGTGAGTGAAG 1080
DB 4525 TGTGACAGCAATATCTGCTGTGACACGTGGTCCGGAAGAGGGGCTGTGAGTGAAG 4584
QY 1081 CGGCTGATAGCGTGTGCTGTGCGGAGTAAACATGTTTCCCGACGCACTATGTCCAG 1140
DB 4585 CGGCTGATAGCGTGTGCTGTGCGGAGTAAACATGTTTCCCGACGCACTATGTCCAG 4644
QY 1141 AGCGAGCGCGACAGTGTCTCATGATCCCTCTCCGACCTTAATACCCGACTTTG 1200
DB 4645 AGCGAGCGCGACAGTGTGTACTAGATCTCTCCAGCCTTAATACCCGACTTTG 4704
QY 1201 AAGAGGCTCCACAGTGAATTAACGAGAGTCTCCACGCGCTGTCCGCTGTGAGCTA 1260
DB 4705 AAGAGGCTCCACAGTGAATTAACGAGAGTCTCTCCAGCCTTAATACCCGACTTTG 4764
QY 1261 AAGGATGTTTGGGACTGATATGACAGTGTGTGCTGACTTGAAGCCTGTGCTGAGTCC 1320
DB 4765 AAGGATGTTTGGGACTGATATGACAGTGTGTGCTGACTTGAAGCCTGTGCTGAGTCC 4824
QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGTCCCAAGTGGATCAAGGG 1380
DB 4825 AAGCTCTGCGCGGAGTAACTGAGTCCCTTTTCTCATGTCCCAAGTGGATCAAGGG 4884
QY 1381 GTCTGGCGGAGAGACGAGTATGACAGACCACTGTCTATGTGAGACAGATACCGGA 1440
DB 4885 GTCTGGCGGAGAGACGAGTATGACAGACCACTGTCTATGTGAGACAGATACCGGA 4944
QY 1441 CATGTCAAAAACGTTCCATGAGATCGTTGGGCTTAAGACCTGTATGAATGTGGCAT 1500
DB 4945 CATGTCAAAAACGTTCCATGAGATCGTTGGGCTTAAGACCTGTGAGCAACAGTGGCAT 5004
QY 1501 GGAATTTCCCATCAACGCAACACACAGGCGCTGTGACCGCTCCCAACGCGCAAC 1560
DB 5005 GGAATTTCCCATCAACGCAACACACAGGCGCTGTGACACCTCTCAAGCGCGCAAC 5064
QY 1561 TAATTCAGGCGCTGTGCGGCTGTGCTGTGAGAGTACGTGAAGTTACGCGGCTGG 1620
DB 5065 TAATTCAGGCGCTGTGCGGCTGTGCTGTGAGAGTACGTGAAGTTACGCGGCTGG 5124
QY 1621 GATTTCCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCGCGTGCAGATTCCA 1680
DB 5125 GATTTCCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCGCGTGCAGATTCCA 5184
QY 1681 GCCCGCGAATTTCTTCAAGAAATGAGTGGGCTGTGCTGACAGATACCTCCGCGTGC 1740
DB 5185 GCTCTGAAATTTCTTCCGAGAGTGAACGAGTGGCGGTTCACAGATACCTCCGCGTGC 5244
QY 1741 AAACCTCTCTACGGAAGAGGTCAATTCAGAGTCCGAGTCAACCAATACCTGTTGG 1800
DB 5245 AAGCTCTCTCTACGGAAGAGGTACATTCAGAGTCCGAGTCAACCAATACCTGTTGG 5304
QY 1801 TCGAGCTTCCATGCGAGGCCAACCAGATGATGAGTCTCACTTCCATGCTCACGAC 1860
DB 5305 TCACAGCTACATGCGAGGCCAACCAGATGATGAGTCTCACTTCCATGCTCACGAC 5364
QY 1861 CCTTCCCATATCAACGACAGACGCTTAAGCCGAGCTGAGCCAGGGGCTCTCCCTTC 1920
DB 5365 CCTTCCCATATCAACGACAGACGCTTAAGCCGAGCTGAGCCAGGGGCTCTCCCTTC 5424
QY 1921 TTGGCAGCTTTCAGCTAGCAGATTGTCTGCGCTTCTCGAAGGCGACATATACCTACC 1980

Db	5425	TTGGCCAGCTCTTCAGCTAGCCAGTTGTCTGCGCCTTCCTTGAGGCGGACATGCACTAAC	5484
Qy	1981	CAAAATGACTTCCAGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGCATGAGATG	2040
Db	5485	CACCATGTCTCTCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGCATGAGATG	5544
Qy	2041	GGC 2043	
Db	5545	GGC 5547	

Search completed: February 25, 2005, 05:54:54
 Job time : 8561 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 01:22:04 ; Search time 957 Seconds
(without alignments)
12637.442 Million cell updates/sec

Title: US-09-664-363-20

Perfect score: 2043

Sequence: 1 TGGGAGGCGCTTCACAGCAG.....TGTGGCGCATGAGATGGGC 2043

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4330206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : N_Geneseq_16Dec04:*

1: geneeqn19808:.*
2: geneeqn19908:.*
3: geneeqn20008:.*
4: geneeqn20018:.*
5: geneeqn20018:.*
6: geneeqn20028:.*
7: geneeqn20028:.*
8: geneeqn20038:.*
9: geneeqn20038:.*
10: geneeqn20038:.*
11: geneeqn20038:.*
12: geneeqn20048:.*
13: geneeqn20048:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2041.4	99.9	3750	AAQ12241	AaQ12241 Encodes P
2	2033.4	99.5	7065	AAQ46195	AaQ46195 PT-NANBH
3	1803	88.3	5955	ACA61114	ACA61114 HCV CDNA
4	1803	88.3	5965	ACA61111	ACA61111 HCV CDNA
5	1803	88.3	37090	ACA61113	ACA61113 Adenovirus
6	1801.4	88.2	7987	ADR38452	Ad38452 DNA encod
7	1801.4	88.2	9587	ADR82189	Ad82189 Hepatitis
8	1796.6	87.9	9416	AAQ20268	AaQ20268 Non-A, no
9	1796.6	87.9	9416	AAQ21829	AaQ21829 Non-A, no
10	1796.6	87.9	9416	AAQ59394	AaQ59394 Non-A, no
11	1787	87.5	7987	AAQ25321	AaQ25321 Hepatitis
12	1787	87.5	7989	AAA98968	AaA98968 Hepatitis
13	1787	87.5	7989	AAQ25322	AaQ25322 Hepatitis
14	1787	87.5	7989	ADJ57845	AdJ57845 HCV repli
15	1787	87.5	7992	AAJ47276	AaJ47276 Hepatitis
16	1787	87.5	8001	AAA98967	AaA98967 Hepatitis
17	1787	87.5	8637	AAA98966	AaA98966 Hepatitis
18	1787	87.5	8638	ABK88595	AbK88595 Hepatitis
19	1787	87.5	8639	ABK88573	AbK88573 Hepatitis
20	1787	87.5	8649	AAA98969	AaA98969 Hepatitis

21	1787	87.5	9605	6	ABK91424	AbK91424 Hepatitis
22	1787	87.5	9605	6	ABK91411	AbK91411 Hepatitis
23	1787	87.5	9605	6	ABK91425	AbK91425 Hepatitis
24	1787	87.5	9605	6	ABK91426	AbK91426 Hepatitis
25	1787	87.5	10690	6	ABK91242	AbK91242 Hepatitis
26	1787	87.5	10690	6	ABK91412	AbK91412 Hepatitis
27	1787	87.5	10690	9	ACA61697	ACA61697 Hepatitis
28	1787	87.5	10690	10	ADC83762	AdC83762 PHCVNeol7
29	1787	87.5	10691	6	ABK91423	AbK91423 Hepatitis
30	1787	87.5	11076	3	AAA98965	AaA98965 Hepatitis
31	1787	87.5	11313	12	ADP86264	AdP86264 Hepatitis
32	1787	87.5	11313	12	ADP86271	AdP86271 Hepatitis
33	1786.6	87.4	8642	6	ABK88574	AbK88574 Hepatitis
34	1785.4	87.4	7141	6	AAQ25333	AaQ25333 Hepatitis
35	1785.4	87.4	7789	6	AAQ25330	AaQ25330 Hepatitis
36	1785.4	87.4	7980	10	ADD93724	AdD93724 Hepatitis
37	1785.4	87.4	7980	10	ADD93725	AdD93725 Hepatitis
38	1785.4	87.4	7987	6	AAQ25329	AaQ25329 Hepatitis
39	1785.4	87.4	7987	6	AAQ25324	AaQ25324 Hepatitis
40	1785.4	87.4	7989	6	AAQ25326	AaQ25326 Hepatitis
41	1785.4	87.4	7989	6	AAQ25325	AaQ25325 Hepatitis
42	1785.4	87.4	7989	10	ADD93726	AdD93726 Hepatitis
43	1785.4	87.4	7989	10	ADD93733	AdD93733 Hepatitis
44	1785.4	87.4	7989	10	ADD93734	AdD93734 Hepatitis
45	1785.4	87.4	7991	6	AAJ47279	AaJ47279 Hepatitis

ALIGNMENTS

RESULT 1	
AAQ12241	
ID	AAQ12241 standard; cDNA; 3750 BP.
XX	AAQ12241;
AC	
XX	
DT	25-MAR-2003 (revised)
DT	17-SEP-1991 (first entry)
XX	
DE	Encodes portion of PT-NANBH viral non-structural protein.
XX	
KW	post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.
XX	
OS	Non-A.
OS	non-B hepatitis virus.
PN	GB2339245-A.
XX	
PD	26-JUN-1991.
XX	
PF	17-DEC-1990; 90GB-00027250.
XX	
PR	18-DEC-1989; 89GB-00028562.
PR	27-FEB-1990; 90GB-00004414.
PR	03-MAR-1990; 90GB-00004814.
XX	
PA	(WELT) WELLCOME FOUND LTD.
PA	(HIGH/) HIGHFIELD P E.
XX	
PI	Highfield PE, Rodgers BC, Tedder RS, Barbara JAJ;
XX	
DR	WPI; 1991-187584/26.
DR	P-PSDB; AAR12599.
XX	
PT	Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA
PT	and antibodies used in diagnostic assays and in vaccines.
XX	
PS	Claim 10; Page 88-97; 108pp; English.
XX	
CC	This sequence probably encodes viral non-structural proteins of the PT-
CC	NANBH viral genome which are antigenic. It was isolated from serum of
CC	humans infected by the virus. See also AAQ12236-40 and AAQ12242. (Updated
CC	on 25-MAR-2003 to correct PA field.)

XX Sequence 3750 BP; 794 A; 1140 C; 1072 G; 744 T; 0 U; 0 Other;
SQ Query Match 99.9%; Score 2041.4; DB 2; Length 3750;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2042; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGGAGGGGCTTTCACAGGCTTCAACCCACGTGATGCCACTTCTGTCACCAACAAG 60
DB 1 TGGAGGGGCTTTCACAGGCTTCAACCCACGTGATGCCACTTCTGTCACCAACAAG 60
QY 61 CAGGAGAGACAATCTTCCCTTACTGCTGGCTTACAGGCTTACTGTGCGCTTAGGGCC 120
DB 61 CAGGAGAGACAATCTTCCCTTACTGCTGGCTTACAGGCTTACTGTGCGCTTAGGGCC 120
QY 121 CAGGCCCCACCTTCATCATGGGATCAAAATGTGAAAGTGTCTCATACGGCTTAAGGCTACT 180
DB 121 CAGGCCCCACCTTCATCATGGGATCAAAATGTGAAAGTGTCTCATACGGCTTAAGGCTACT 180
QY 181 CTGCGCGGGCCAAACCCCTTGTGCTGTATAGGCTGGAGCCGTCACAAACGAGGTCAACCTTC 240
DB 181 CTGCGCGGGCCAAACCCCTTGTGCTGTATAGGCTGGAGCCGTCACAAACGAGGTCAACCTTC 240
QY 241 ACACACCCCATTAACCAATTCATCATGCGCATGATGTCAAGCCGACTTGGAGGTCTGCACG 300
DB 241 ACACACCCCATTAACCAATTCATCATGCGCATGATGTCAAGCCGACTTGGAGGTCTGCACG 300
QY 301 AGCACCTGGGTGTGTGGTGGCGGGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTT 360
DB 301 AGCACCTGGGTGTGTGGTGGCGGGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTT 360
QY 361 GGCAGCGGTGTGTGTGGTGTGAGTATCATCTTGTCCGGCGGCGGCTATTGTTCCCGAC 420
DB 361 GGCAGCGGTGTGTGTGGTGTGAGTATCATCTTGTCCGGCGGCGGCTATTGTTCCCGAC 420
QY 421 AGGGAAGTCTTACCAAGAGTTCCATGATGAGTGAAGAGTGGCGCTTCCGACTTCCCTTAC 480
DB 421 AGGGAAGTCTTACCAAGAGTTCCATGATGAGTGAAGAGTGGCGCTTCCGACTTCCCTTAC 480
QY 481 ATCGAGAGGGAATGAGCTCGCCGAGCAGTTCAAGCAAAAAGCGCTCGGGTGTGTCAG 540
DB 481 ATCGAGAGGGAATGAGCTCGCCGAGCAGTTCAAGCAAAAAGCGCTCGGGTGTGTCAG 540
QY 541 ACAGCCACCACCAAGCGAGGCGCGTGTCTCCCGTGTGTGAGTCCAAATGTGCGAGCCCTT 600
DB 541 ACAGCCACCACCAAGCGAGGCGCGTGTCTCCCGTGTGTGAGTCCAAATGTGCGAGCCCTT 600
QY 601 GAGACCTTCTGGGCGCAACACATGTGGAATTTCATCAGCGGGAATACAGTACTTAGCAGGC 660
DB 601 GAGACCTTCTGGGCGCAACACATGTGGAATTTCATCAGCGGGAATACAGTACTTAGCAGGC 660
QY 661 TTGTCCACTCTGCTGGGAATCCCGGATGTGATCATCTAGTGGCGTTCAAGCTCTGTC 720
DB 661 TTGTCCACTCTGCTGGGAATCCCGGATGTGATCATCTAGTGGCGTTCAAGCTCTGTC 720
QY 721 ACTAGCCGCTCAACCAATTCATCTCTGCTTAAATCATCTGGGGGATGAGTGAAGC 780
DB 721 ACTAGCCGCTCAACCAATTCATCTCTGCTTAAATCATCTGGGGGATGAGTGAAGC 780
QY 781 GCCCACTGCTCCCCCAGTGTGCTTCAAGCTTCTGTAAGGCGCGGCAATGCTGTGTCG 840
DB 781 GCCCACTGCTCCCCCAGTGTGCTTCAAGCTTCTGTAAGGCGCGGCAATGCTGTGTCG 840
QY 841 GGTGTTGGAGATAGAGCTTGGGAAGTGTGTTGGAATCATCTGGCGGGATATGAGAGA 900
DB 841 GGTGTTGGAGATAGAGCTTGGGAAGTGTGTTGGAATCATCTGGCGGGATATGAGAGA 900
QY 901 GGAAGTGGCAGGCGCTGCTGAGCTTTAAAGTCAATAGCGCGAAATGCTTCAACCGAG 960
DB 901 GGAAGTGGCAGGCGCTGCTGAGCTTTAAAGTCAATAGCGCGAAATGCTTCAACCGAG 960
QY 961 GACTGTGTTAACTTACTCTCTGCGATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 GACTGTGTTAACTTACTCTCTGCGATCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

DB 961 GACTGTGTTAACTTACTCTCTGCGATCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 TGGCAGGAGATATCTGCTGCGACGCTGGGTCCAGGGAGGGGGCTGTCACTGATGATGAAC 1080
DB 1021 TGGCAGGAGATATCTGCTGCGACGCTGGGTCCAGGGAGGGGGCTGTCACTGATGATGAAC 1080
QY 1081 CGGCTGATAGGTTTGCTTCCGGGGGTAACCATGTTTCCCCCAGCCACTATGTGCAAGAG 1140
DB 1081 CGGCTGATAGGTTTGCTTCCGGGGGTAACCATGTTTCCCCCAGCCACTATGTGCAAGAG 1140
QY 1141 AGCGAGCGCGAGACAGTGTCACTGATATCCCTTCCGACTTATCATGACCCAACTGTG 1200
DB 1141 AGCGAGCGCGAGACAGTGTCACTGATATCCCTTCCGACTTATCATGACCCAACTGTG 1200
QY 1201 AAGAGCTTCAACCAATGATTAACAGAGACTGCTTCAAGCCCTGCTCCGGCTGAGCTTA 1260
DB 1201 AAGAGCTTCAACCAATGATTAACAGAGACTGCTTCAAGCCCTGCTCCGGCTGAGCTTA 1260
QY 1261 AGGATGTTTGGAGCTGATATGACACAGTTTGGCTGACTTCAAGACTGCTCCAGTTC 1320
DB 1261 AGGATGTTTGGAGCTGATATGACACAGTTTGGCTGACTTCAAGACTGCTCCAGTTC 1320
QY 1321 AAGCTCTGCGCGGATTAACCGGGAGTCCCTTTTCTCATGCGCAAGTGGTAAAGGGG 1380
DB 1321 AAGCTCTGCGCGGATTAACCGGGAGTCCCTTTTCTCATGCGCAAGTGGTAAAGGGG 1380
QY 1381 GTCTGGCGGGAGACGCGATCATGACAGACCACTGCTCATGTGAGACAAGATCACCGGA 1440
DB 1381 GTCTGGCGGGAGACGCGATCATGACAGACCACTGCTCATGTGAGACAAGATCACCGGA 1440
QY 1441 CATGTCAAAAACGGTTCATGAGATCGTTGGGCTTAAGACTGTATGTAACATGTGGCAT 1500
DB 1441 CATGTCAAAAACGGTTCATGAGATCGTTGGGCTTAAGACTGTATGTAACATGTGGCAT 1500
QY 1501 GGAACATTCGCCATCAAGCATACACACAGCGGCCCTTGCAGCGCTCCCGACGCGCAAC 1560
DB 1501 GGAACATTCGCCATCAAGCATACACACAGCGGCCCTTGCAGCGCTCCCGACGCGCAAC 1560
QY 1561 TATTCAGGCGCTGTGGCGGGTGGCTGTGAGAGTACGTGAGGTTACGCGGGTGGGG 1620
DB 1561 TATTCAGGCGCTGTGGCGGGTGGCTGTGAGAGTACGTGAGGTTACGCGGGTGGGG 1620
QY 1621 GATTTTCACATAGTACAGACATGACCACTGACCAACGTAATAATGCCCGTCCAGGTTCCA 1680
DB 1621 GATTTTCACATAGTACAGACATGACCACTGACCAACGTAATAATGCCCGTCCAGGTTCCA 1680
QY 1681 GCCCGCGAATCTTCAAGAAATGATGGAGTGGCGGCTGACAGGTAAGCTCCGCGGTTC 1740
DB 1681 GCCCGCGAATCTTCAAGAAATGATGGAGTGGCGGCTGACAGGTAAGCTCCGCGGTTC 1740
QY 1741 AAACCTCTCTTACGGAAGAGTCAATTCACAGTCCGAGCTTCACCAATCACTGTTGGG 1800
DB 1741 AAACCTCTCTTACGGAAGAGTCAATTCACAGTCCGAGCTTCACCAATCACTGTTGGG 1800
QY 1801 TCGCAGCTTCCATGCGAGCGGACCGAATGTAAGTGAAGTCTCACTTCAATGCTCACCGAC 1860
DB 1801 TCGCAGCTTCCATGCGAGCGGACCGAATGTAAGTGAAGTCTCACTTCAATGCTCACCGAC 1860
QY 1861 CCTTCCCATATACAGCAGAGCGGCTAAGCCAGAGCTGAGCCAGGGGGTCTCCCGCTTC 1920
DB 1861 CCTTCCCATATACAGCAGAGCGGCTAAGCCAGAGCTGAGCCAGGGGGTCTCCCGCTTC 1920
QY 1921 TTGGCCAGCTCTTCACTAGCAGTGTGTGCGCTTCTCGAAGCGCATATCATATACC 1980
DB 1921 TTGGCCAGCTCTTCACTAGCAGTGTGTGCGCTTCTCGAAGCGCATATCATATACC 1980
QY 1981 CAAAATGACTTCCAGACGCTGACTCATGAGGCGCAACCTCTGTGGCGGATGAGATG 2040
DB 1981 CAAAATGACTTCCAGACGCTGACTCATGAGGCGCAACCTCTGTGGCGGATGAGATG 2040
QY 2041 GGC 2043
DB 2041 GGC 2043


```

Db      3484 AGGAGTGTGGAGCTGATATGACAGTTTTGGCTGACTTCAGAACCTGGCTCCAGTTC 3543
Oy      1321 AAGCTCTGCGCGATTAACCGGGAGTCCCTTTTCTCATGCGCAAGTGGATCAAGGGG 1380
Db      3544 AAGCTCTGCGCGATTAACCGGGAGTCCCTTTTCTCATGCGCAAGTGGATCAAGGGG 3603
Oy      1381 GTCTGGCGGGAGACGGCATCATGACACCACTTGCTCATGTGGACACAGTACCGGA 1440
Db      3604 GTCTGGCGGGAGACGGCATCATGACACCACTTGCTCATGTGGACACAGTACCGGA 3663
Oy      1441 CATGCAAAAAGGTTCCATGAGATCGTTGGGCTTAAGACCTGATAGTAACTAGTGGCAT 1500
Db      3664 CATGCAAAAAGGTTCCATGAGATCGTTGGGCTTAAGACCTGATAGTAACTAGTGGCAT 3723
Oy      1501 GGAATATTCCTCCATACGATACACCAACGAGGCGCCCTGACCGCTCCCGACGCGCAAC 1560
Db      3724 GGAATATTCCTCCATACGATACACCAACGAGGCGCCCTGACCGCTCCCGACGCGCAAC 3783
Oy      1561 TATTCAGGGCGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGATTACGCGGGTGGGG 1620
Db      3784 TATTCAGGGCGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGATTACGCGGGTGGGG 3843
Oy      1621 GATTTTCCATAGTACGACGACGATGACCACTGACCAAGTAAATGCCCGCGCAGGTTCGA 1680
Db      3844 GATTTTCCATAGTACGACGACGATGACCACTGACCAAGTAAATGCCCGCGCAGGTTCGA 3903
Oy      1681 GCCCCCGAATTTCTTCAAGAAATGATGAGGGGTGGCGGCTGACAGGATCGCTCCGGCGTTC 1740
Db      3904 GCCCCCGAATTTCTTCAAGAAATGATGAGGGGTGGCGGCTGACAGGATCGCTCCGGCGTTC 3963
Oy      1741 AAACCTTCTCTTACGAGAGAGGTTCACATTCACAGTTCGGGCTCAACCAATATCTGTTGGG 1800
Db      3964 AAACCTTCTCTTACGAGAGAGGTTCACATTCACAGTTCGGGCTCAACCAATATCTGTTGGG 4023
Oy      1801 TGGGAGTCTCCCATGGAAGCGCGCAACCGGATGATGACAGTCTCACTTCCATGCTCACCGAC 1860
Db      4024 TGGGAGTCTCCCATGGAAGCGCGCAACCGGATGATGACAGTCTCACTTCCATGCTCACCGAC 4083
Oy      1861 CCTCCCATCATACAGACAGACGCGTACCGGCTAGCCAGGCTGGCGAGGGGGTCTCCCGCTTC 1920
Db      4084 CCTCCCATCATACAGACAGACGCGTACCGGCTAGCCAGGCTGGCGAGGGGGTCTCCCGCTTC 4143
Oy      1921 TTGGCCAGCTTTCAGCTAGTACCAAGTGTCTGGCGCTTCTCGAAGGCGACATATATAC 1980
Db      4144 TTGGCCAGCTTTCAGCTAGTACCAAGTGTCTGGCGCTTCTCGAAGGCGACATATATAC 4203
Oy      1981 CAAATGATTTCCCGACGAGTACCTCATCGAGGCGCAACCTCTGTTGGCGGATGAGATG 2040
Db      4204 CAAATGATTTCCCGACGAGTACCTCATCGAGGCGCAACCTCTGTTGGCGGATGAGATG 4263
Oy      2041 GGC 2043
Db      4264 GGC 4266

RESULT 3
ACAE1114
ID      ACAE1114 strand; cDNA; 5955 BP.
XX
AC      ACAE1114;
XX
DT      09-JUL-2003 (first entry)
XX
DE      HCV cDNA encoding Met-NS3-NS4A-NS4B-NS5A-NS5B (active).
XX
XX      HCV; ser; gene; non-structural protein; NS3; NS4A; NS4B; NS5A; NS5B;
XX      adenoviral vector; HCV infection; vaccine; gene therapy; protease.
XX
OS      Hepatitis C virus.
XX
FH      Key Location/Qualifiers
FT      CDS 1..5955
FT      /tag= a

```

```

FT      /product= "Met-NS3-NS4A-NS4B-NS5A-NS5B"
FT      /partial
FT      /note= "No stop codon shown"
XX
XX      W02003031588-A2.
XX
XX      17-APR-2003.
XX
XX      10-OCT-2002; 2002WO-US032512.
XX
XX      11-OCT-2001; 2001US-0328655P.
XX      13-MAR-2002; 2002US-0363774P.
XX
XX      (MERI ) MERCK & CO INC.
XX      (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX      Emili BA, Kariow DC, Bett AJ, Shiver JW, Nicosia A, Lahm A;
XX      Luzzago A, Cortese R, Colloca S;
XX      WPI; 2003-381708/36.
XX      P-PsDB; ABU09575.
XX
XX      New nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A-NS5B polypeptide,
XX      useful as a component of an adenovector or DNA plasmid vaccine for
XX      preventing or treating hepatitis C virus.
XX
XX      Example 1; Fig 5; 231pp; English.
XX
XX      The invention relates to a nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A
XX      -NS5B (NS stands for non-structural protein from HCV) polypeptide
XX      appearing as ABU09574. The encoded polypeptide has sufficient protease
XX      activity to process itself to produce an NS5B protein that is
XX      enzymatically inactive. Also included are a cultured recombinant cell
XX      comprising the novel nucleic acid, making an adenovector (comprising: (a)
XX      producing an adenovirus genome plasmid comprising a gene expression
XX      cassette by homologous recombination between the novel nucleic acid and a
XX      nucleic acid comprising a first adenovirus region from base pair 1-450
XX      corresponding to either Ad5 or Ad6, a second adenovirus region from base
XX      pair 3511-5548 corresponding to Ad5 or from base pair 3508-5541
XX      corresponding to Ad6, joined to the first region, a third adenovirus
XX      region from base pair 5549-2813 corresponding to Ad5 or from base pair
XX      5542-28156 corresponding to Ad6, joined to the second region, a fourth
XX      adenovirus region from base pair 30818-33966 corresponding to Ad5 or from
XX      base pair 30789-33784 corresponding to Ad6, joined to the third region,
XX      and a fifth adenovirus region from base pair 33967-35935 corresponding to
XX      Ad5 or from base pair 33785-35759 corresponding to Ad6, joined to the
XX      fourth region; and (b) rescuing the recombinant adenovirus from the
XX      recombinant adenovirus plasmid, an adenoviral vector that is produced by
XX      method above appearing as ACAE1113 which has a humanised version of the
XX      polynucleotide of the invention and encodes the HCV inactivated
XX      polynucleotide, a recombinant nucleic acid comprising one or more Ad6
XX      regions and a region not present in Ad6, where at least one Ad6 region is
XX      selected from E1A, E1B, E2A, E2B, E4, L1, L2, L4 and L5. The nucleic acid
XX      is useful as a component of an adenoviral vector or a DNA plasmid vaccine
XX      providing a broad range of antigens for generating an HCV-specific cell
XX      mediated immune response. The nucleic acid may also be used in treating
XX      patients infected with HCV. The present sequence is the non-humanised
XX      cDNA encoding the wild-type HCV polypeptide with an active NS5B protein
XX
XX      Sequence 5955 BP; 1251 A; 1798 C; 1678 G; 1228 T; 0 U; 0 Other;

```

```

Query Match      88.3%; Score 1803; DB 8; Length 5955;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

```

```

Oy      1 TGGAGGGCGCTTTTCAACAGGCTTCAACCGATGATGCCACTTCTGCGCAACCAAG 60
XX      |||||
XX      1597 TGGAGAGTGTCTTCAAGGCTTCAACCAATAGATGACACTTCTTGTCCAGACCAAG 1656
Db      1597 TGGAGAGTGTCTTCAAGGCTTCAACCAATAGATGACACTTCTTGTCCAGACCAAG 1656
Oy      61 CAGGAGAGACCAACTCCCTACTGTTGGGCTACAGGCTACTGTGTGCGCTAGGGCC 120
XX      |||||
XX      1657 CAGGAGAGACCAACTCCCTACTGTTGGGCTACAGGCTACTGTGTGCGCTAGGGCC 1716
Db      1657 CAGGAGAGACCAACTCCCTACTGTTGGGCTACAGGCTACTGTGTGCGCTAGGGCC 1716

```

QY 121 CAGGCCCACTTCATCATGAGATCAAAATGAGAGTGTCTCATACGAGCTAAAGCTACT 180
 Db 1717 CAGGCCCACTTCATCATGAGATCAAAATGAGAGTGTCTCATACGAGCTAAAGCTACT 1776
 QY 181 CTGCGCGGCGCAACACCTTCTGTGTATAGGCTGGAGCGCTCCAAAACGAGTCAACCTTC 240
 Db 1777 CTGCAACGGGCGCAACACCTTCTGTGTATAGGCTGGAGCGCTCCAAAACGAGTCAACCTTC 1836
 QY 241 ACACACCCCATTAACAAATTCATCATGAGATCAATGATGACGAGACCTGGAGGTCGTACAG 300
 Db 1837 ACCCACCCCATTAACAAATTCATCATGAGATCAATGATGACGAGACCTGGAGGTCGTACACT 1896
 QY 301 AGCAGCTGGAGTCTGGTGGCGGAGTCTTTCAGAGCTCTGAGTGGATGCTTGAACAGA 360
 Db 1897 AGCAGCTGGAGTCTGGTGGCGGAGTCTTTCAGAGCTCTGAGTGGATGCTTGAACAGA 1956
 QY 361 GGCAGCGTGTATGTTGGGTAGATCATCTTGTCCGGCGGCGGCTATTTGTTCCCGAC 420
 Db 1957 GGCAGTGTGTATGTTGGGTAGATCATCTTGTCCGGCGGCGGCTATTTGTTCCCGAC 2016
 QY 421 AGGGAAGTCTCTACAGAGATTCATGATGAGAGAGTGGGCGTGGACCTCCCTTAC 480
 Db 2017 AGGGAAGTCTCTACAGAGATTCATGATGAGAGAGTGGGCGTGGACCTCCCTTAC 2076
 QY 481 ATCGACAGGGAATGAGCTCGCGGAGAGTTCAGACAAAACGCTCGGTTGCTGCAG 540
 Db 2077 ATCGAGAGGGAATGAGCTCGCGGAGAGTTCAGACAAAACGCTCGGTTGCTGCAG 2136
 QY 541 ACAGCCACCAAGCAAGCGGAGCGCTGCTCCGCTGTGTGAGTCCAAATGGCGAGCCCTT 600
 Db 2137 ACAGCCACCAAGCAAGCGGAGCGCTGCTCCGCTGTGTGAGTCCAAATGGCGAGCCCTT 2196
 QY 601 GAGACCTTCTGGGCGAAACATGTTGAACTTCATACAGGAGATTAAGTACTTAAGAGAGC 660
 Db 2197 GAGACCTTCTGGGCGAAACATGTTGAACTTCATACAGGAGATTAAGTACTTAAGAGAGC 2256
 QY 661 TTGTCACCTCTGCTGGGAATCCCGGATGATCATCTAGTGGCGCTTCAAGCTCTGTC 720
 Db 2257 TTATCATCTCTGCTGGGAATCCCGGATGATCATCTAGTGGCGCTTCAAGCTCTGTC 2316
 QY 721 ACTAGCCGCTCACCAACCAATCTACCTCTGCTTAACTCTGGGAGGATAGAGCC 780
 Db 2317 ACTAGCCGCTCACCAACCAATCTACCTCTGCTTAACTCTGGGAGGATAGAGCC 2376
 QY 781 GCCCAATCTGCTCCCCCGAGTCTGTTAGCTTTGTAGGGCGCGCATTTGTTGGTGG 840
 Db 2377 GCCCAATCTGCTCCCCCGAGTCTGTTAGCTTTGTAGGGCGCGCATTTGTTGGTGG 2436
 QY 841 GCTGTTGGCAGATAGGCTTGGGAAGTGTGTGAGATCTTGGCGGAGTATGAGAGA 900
 Db 2437 GCTGTTGGCAGATAGGCTTGGGAAGTGTGTGAGATCTTGGCGGAGTATGAGAGA 2496
 QY 901 GGAAGTGGAGGCGCGCTGCTGCTTTAAGTATAGAGCGGAGAAATGCTTCCACGAG 960
 Db 2497 GGAAGTGGAGGCGCGCTGCTGCTTTAAGTATAGAGCGGAGAAATGCTTCCACGAG 2556
 QY 961 GACCTGTTAACTTACTCTCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db 2557 GACCTGTTAACTTACTCTCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2616
 QY 1021 TGCAGAGCATAGTGTGAGAGCGTGGGTCCAGGGAGGAGGCTGTGAGAGTGAAG 1080
 Db 2617 TGCAGAGCATAGTGTGAGAGCGTGGGTCCAGGGAGGAGGCTGTGAGAGTGAAG 2676
 QY 1081 CGGCTGATAGCGTTCGCTCGCGGGGTAAACATGTTTCCCGACGACATATGTGCCAG 1140
 Db 2677 CGGCTGATAGCGTTCGCTCGCGGGGTAAACATGTTTCCCGACGACATATGTGCCAG 2736
 QY 1141 AGCGAGCGCGGAGAGTGTCTAGATCTCTCCGAGCTTACTATACCCCACTGTTG 1200
 Db 2737 AGCGAGCGCGGAGAGTGTCTAGATCTCTCTCCGAGCTTACTATACCCCACTGTTG 2796
 QY 1201 MAGAGCTCCACAGTGAATTAACAGAGACTGTCTCAACGCCCTGCTCGGCTGTGGCTA 1260

Db 2797 AAAAGGCTCCACAGTGAATTAAGAACTGCTCCACACCGTGTCCCGGCTGTGGCTA 2856
 QY 1261 AGGAGTGTGGGACCTGATTAAGCAAGTTTGGCTGACTTGAAGCTGGGCTCAAGTCC 1320
 Db 2857 AGGAGTGTGGGACCTGATTAAGCAAGTGTGACTGACTTGAAGCTGGGCTCAAGTCC 2916
 QY 1321 AAGCTCTCCCGGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTCAAGAGG 1380
 Db 2917 AAGCTCTCCCGGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTCAAGAGG 2976
 QY 1381 GTCTGGGAGGAGAGCGGATTCATGACACACCTGCTCATGTGAGACAGATCACCGGA 1440
 Db 2977 GTCTGGGAGGAGAGCGGATTCATGACACACCTGCTCATGTGAGACAGATCACCGGA 3036
 QY 1441 CATGTCAAAAAGGTTCCATGAGATCTGTGGGCTTAACCTGTAGTAACTATGGGCA 1500
 Db 3037 CATGTCAAAAAGGTTCCATGAGATCTGTGGGCTTAACCTGTAGTAACTATGGGCA 3096
 QY 1501 GGAACATTCGCCATCAACGATACACACGAGGCGCTGACGCGCTCCCAAGCGCAAC 1560
 Db 3097 GGAACATTCGCCATCAACGATACACACGAGGCGCTGACGCGCTCCCAAGCGCAAC 3156
 QY 1561 TATTCAGGCGCTGTGTGCGGAGTGTGCTGAGAGTACGTGAGGTTACGCGGTTGG 1620
 Db 3157 TATTCAGGCGCTGTGTGCGGAGTGTGCTGAGAGTACGTGAGGTTACGCGGTTGG 3216
 QY 1621 GATTTCCATAGTATGACAGGATGACCACTGACAAAGTAAATGCCCGGCTCAAGTTC 1680
 Db 3217 GATTTCCATAGTATGACAGGATGACCACTGACAAAGTAAATGCCCGGCTCAAGTTC 3276
 QY 1681 GCGCCGGAATTCCTCAAGAGTGAATGGGATGGGCTGACAGGATGCTCGGAGTGC 1740
 Db 3277 GCTCTGAATTCCTCAAGAGTGAATGGGATGGGCTGACAGGATGCTCGGAGTGC 3336
 QY 1741 AAACTCTCTCAAGGAGAGGTCAATTCAGTGTGGGCTCAACCAATACCTGTTGG 1800
 Db 3337 AGGCTCTCTCAAGGAGAGGTTCATTCAGGATGGGCTCAACCAATACCTGTTGG 3396
 QY 1801 TGCAGCTCCATGAGAGCGCGGAAACGATGTAAGAGTGTCACTTCCATGCTCACCGAC 1860
 Db 3397 TGCAGCTCCATGAGAGCGCGGAAACGATGTAAGAGTGTCACTTCCATGCTCACCGAC 3456
 QY 1861 CCTCCACATCAACAGCAGAGAGCGGTAAGCGGAGGCTGAGGAGGCTCCCGCTCC 1920
 Db 3457 CCTCCACATCAACAGCAGAGAGCGGTAAGCGGAGGCTGAGGAGGCTCCCGCTCC 3516
 QY 1921 TTGGCAGCTCTTCAAGCTAGCCAGTGTGCGCTTCTCTGAAAGGCGCATATCATTTAC 1980
 Db 3517 TTGGCAGCTCTTCAAGCTAGCCAGTGTGCGCTTCTCTGAAAGGCGCATATCATTTAC 3576
 QY 1981 CAAATGACTTCCAGAGCGGTGACTCATGAGGCGCAACCTCTGTGGCGGATGAGAG 2040
 Db 3577 CAAATGACTTCCAGAGCGGTGACTCATGAGGCGCAACCTCTGTGGCGGAGTGA 3636
 QY 2041 GGC 2043
 Db 3637 GGC 3639

RESULT 4
 ID ACA61111 standard; cDNA; 5965 BP.
 XX ACA61111;
 AC XX
 AC XX
 DT 09-JUL-2003 (first entry)
 XX
 DE HCV cDNA encoding Met-NS3-NS4A-NS4B-NS5A-NS5B (inactivated).
 HCV; ss; gene; non-structural protein; NS3; NS4A; NS4B; NS5A; NS5B;
 KM adenoviral vector; HCV infection; vaccine; mutant; gene therapy;
 protease.

Query Match	Best Local Similarity	Score 1003;	DB 8;	Length 5965;
Matches 1003;	Conservative	0;	Mismatches 150;	Indels 0;
Gaps	0			

QY	I	TGGAGGGCGCTTTTCA	CAGGCGCTTCA	CCCA	CGTGGATGCC	CACTTCTGTCTCCCAACAAG	60	
Db	1603	TGGAGAGTGTCTTTC	CAAGAGCTTCA	CCCA	TATATCA	CACTTCTGTCTCCAGACAAG	1662	
QY	61	CAGGCAAGAGCAACTT	CCCCCTAC	CTGTGGCCGTA	CCAGGCTAC	TGTGTGCGTAGAGGCC	120	
Db	1663	CAGGCAAGAGCAACTT	CCCCCTAC	CTGTGTAGCAT	CCAA	CGCGTGTGCGCAGGGCT	1722	
QY	121	CAGGCCCCACCTCCAT	CATGGGATCA	AAATGTGAA	GTGTCTA	TCGGCTTAAAGCTTACT	180	
Db	1723	CAGGCCCCACCTCCAT	CATGGGATCA	AAATGTGAA	GTGTCTA	TCGGCTTAAAGCTTACT	1782	
QY	181	CTGCGCGGGCCAC	ACCTTGTCTGT	ATA	GGCTGGAGAC	CGCTCCAAAACAGAGTCA	240	
Db	1783	CTGCACGGGGCCAA	CACTTGTCTGT	ATA	CGGCTGGAGAC	CGCTCCAAAATGAGGTCA	1842	
QY	241	ACACACCCCAATPAC	CAAAATTCAT	CATATGAG	CAATG	TCAGCCGACCTGGAGAGTCTGACG	300	
Db	1843	ACCAACCCCATACCA	AAATACATATG	CAATG	CAATG	TCAGCTGACCTGGAGAGTCTGACT	1902	
QY	301	AGCACTTGGGTGTGT	GTGGCGGGGTCTT	GAC	ACTTGGCTGTG	GTATTTGCTTGACACA	360	
Db	1903	AGCACTTGGGTGTGT	GTGGCGGGGTCTT	GAC	ACTTGGCTGTG	GTATTTGCTTGACACA	1962	
QY	361	GGCAGCGTGTCA	TTTGTGGGTAG	AGATCAT	TTTGTCCGGCGGGCCGCTAT	TTTGTCTCCGAC	420	
Db	1963	GGCAGTGTGTCA	TTTGTGGGTAG	AGATTA	TCCTTGTCCGGAGGGCCGGCTAT	TTTGTCTCCGAC	2022	
QY	421	AGGAAAGTCCCTA	TCACAGAGATTG	AGATGAG	ATGAG	TGGAAAGGCGCGTGGCACTCCCTTAC	480	
Db	2023	AGGAAATTTCTTAC	CAAGAAATTCAT	GA	TAAATGGA	AGATGCGCTCTGCACTTCCCTTAC	2082	
QY	481	ATTCAGACAGGAA	TGACAGCTTCC	CGACAGTTC	CAAGCAAAAAGCGCT	TCGGTTGTCTGACG	540	
Db	2083	ATTCAGACAGGAA	TGACAGCTTCC	CGACACA	ATTCAAGCAAAAAGCGCT	TCGGTTTACTGCAA	2142	
QY	541	ACAGCCACCAAG	CAAGCGAGGCGCTG	CTCCGTGTG	AGATCTCA	AGTGGCGAGCCCTT	600	
Db	2143	ACAGCCACCAAG	CAAGCGAGGCTG	CTCCGTGTG	AGATCTCA	AGTGGCGAGCCCTT	2202	
QY	601	GAGACCTTCTGGG	GGGAAACA	TGTGGA	CTTCA	TACGCGGGAATACAGTACTTATGACAGGC	660	
Db	2203	GAGACCTTCTGGG	GGGAAACA	TGTGGA	CTTCA	TACGCGGGAATACAGTACTTATGACAGGC	2262	
QY	661	TTGTCACTCTG	CGTGGGAATCC	CGCATTTG	GCATCA	CTGTATGAGCGTTCA	720	
Db	2263	TTATTCAC	TCTGCTGGGA	ATCCCGCA	ATAGCAT	ATGATGCA	720	
QY	721	ACTAGCCCGCTAC	CAACCCATTA	CCCTCTG	CTTAA	CACTCTGGGGGGAATGGGTATGC	780	
Db	2323	ACCAAGCCCGCTAC	CAACCCATTA	CCCTCTG	CTTAA	CACTCTGGGGGGAATGGGTATGC	2382	
QY	781	GCCCAACTCGCT	CCCCCGCATG	CGCTTCA	GGCTTTC	GAAGGCGCGGCAATGCTGTGCG	840	
Db	2383	GCCCAACTCGCT	CCCCCGCATG	CGCTTTC	GAAGGCGCGGCAATGCTGTGCG	2442		
QY	841	GCTGTGGACAT	ATAGGCTTTGG	GAAGGTGCTTTG	ACATCTT	TGAGCGCGGCTATGAGCA	900	
Db	2443	GCTGTGGACAT	ATAGGCTTTGG	GAAGGTGCTTTG	ACATCTT	TGAGCGCGGCTATGAGCA	2502	
QY	901	GGAAGTGGACGG	CGCGCTGTGG	CGCTTTA	AGATCAT	ATAGCGGCGGAATGCTTCCACCGAG	960	
Db	2503	GGAAGTGGACGG	CGCGCTGTGG	CGCTTTA	AGATCAT	ATAGCGGCGGAATGCTTCCACCGAG	2562	
QY	961	GACCTGTGTAA	CTTA	CTCCCTG	GCATCTCTCTG	TGCGCGCTGTGTGCTGTGGGTGCTG	1020	
Db	2563	GACCTGTGTAA	CTTA	CTCCCTG	GCATCTCTCTG	TGCGCGCTGTGTGCTGTGGGTGCTG	2622	
QY	1021	TGGCAGAGAT	ACTAGCGCTCG	CA	GTGGGTTC	CAAGGAGGGGCGCTGTGCA	1080	
Db	2623	TGTGACAGAA	TA	CTGTGCTG	CA	AGTGGGTTC	CGGGAAGGGGCGCTGTGCA	2682
QY	1081	CGGCTGAT	AGGTTTGC	CTTGC	CGGGGGTAA	CCATGTTTCTCCCA	CGCACTATGTGCGAG	1140


```

Db      ||||| 2683 CCGCGATGAGCTTGGCCCTCGCGGGGTATCATGTTCCGCCACGCACTATGCTGAG 2742
Oy      ||||| 1141 AGCGAGCGCGGAGACGCTGCTCACTAGATCCCTCCGACTTACTATCAACCACTGTTG 1200
Db      ||||| 2743 AGCGAGCGCGGAGACGCTGCTCACTAGATCCCTCCGACTTACTATCAACCACTGTTG 2802
Oy      ||||| 1201 AAGAGGCTCCACGAGTGAATTAACGAGAGCTGCTCCAGCCCTGCTCCGGCTGTTGCTGA 1260
Db      ||||| 2803 AAAAGGCTCCACGAGTGAATTAATGAGAGCTGCTCCAGCCCTGCTCCGGCTGTTGCTGA 2862
Oy      ||||| 1261 AGGAGTGTGGGAGCTGATATGACAGATTTTGGCTGACTTCAAGACTGGCTCCAGTCC 1320
Db      ||||| 2863 AGGAGTGTGGGAGCTGATATGACAGATTTTGGCTGACTTCAAGACTGGCTCCAGTCC 2922
Oy      ||||| 1321 AAGCTTCGCGCGGAGTATACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAGAGGG 1380
Db      ||||| 2923 AAGCTTCGCGCGGAGTATACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAGAGGG 2982
Oy      ||||| 1381 GTCTGGCGGGAGACGAGCATCATGACGACCACTGCTCATGTGAGACAGATCAACGGA 1440
Db      ||||| 2983 GTCTGGCGGGAGACGAGCATCATGACGACCACTGCTCATGTGAGACAGATCAACGGA 3042
Oy      ||||| 1441 CATGTCAAAAACGGTTCATGAGAGATGTTGGGCTTAAGACCTGTATGATGATGTCAT 1500
Db      ||||| 3043 CATGTCAAAAACGGTTCATGAGAGATGTTGGGCTTAAGACCTGTGACGACCAAGTGGCAT 3102
Oy      ||||| 1501 GGAACATTTCCCATCAAGCATACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
Db      ||||| 3103 GGAACATTTCCCATCAAGCATACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3162
Oy      ||||| 1561 TATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Db      ||||| 3163 TATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3222
Oy      ||||| 1621 GATTTCCACTACGTGACGAGCATGACACTGACCAAGTAAATGCCCCTGACAGGTTCCA 1680
Db      ||||| 3223 GATTTCCACTACGTGACGAGCATGACACTGACCAAGTAAATGCCCCTGACAGGTTCCA 3282
Oy      ||||| 1681 GCGCCCGAATTTCTTCAAGAAATGAGATGGAGGTGCGGCTGCAAGATGAGTCCGGGCTGC 1740
Db      ||||| 3283 GCTCTGAATTTCTTCAAGAAATGAGATGGAGGTGCGGCTGCAAGATGAGTCCGGGCTGC 3342
Oy      ||||| 1741 AAACCTCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
Db      ||||| 3343 AGGCTCTCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3402
Oy      ||||| 1801 TCGAGAGCTTCCATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
Db      ||||| 3403 TCACAGCTACATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3462
Oy      ||||| 1861 CCGTCCCATCATCAAGCAGAGACGCTTAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Db      ||||| 3463 CCGTCCCATCATCAAGCAGAGACGCTTAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3522
Oy      ||||| 1921 TTGGCAGAGCTTTCAGTACGATGATGTTGGGCTTCTCGAAGGCGCATTAACATTACC 1980
Db      ||||| 3523 TTGGCAGAGCTTTCAGTACGATGATGTTGGGCTTCTCGAAGGCGCATTAACATTACC 3582
Oy      ||||| 1981 CAAAATGATTTCCAGAGCGCTGACCTCATCAAGGCGCAACCTCTGTGGCGGATGAGATG 2040
Db      ||||| 3583 CACCATGATTTCCAGAGCGCTGACCTCATCAAGGCGCAACCTCTGTGGCGGATGAGATG 3642
Oy      ||||| 2041 GGC 2043
Db      ||||| 3643 GGC 3645

```

RESULT 5
 ACA61113
 ID ACA61113 standard; DNA; 37090 BP.
 XX
 AC ACA61113;

```

XX      09-JUL-2003 (first entry)
DT      Adenovirus vector MRKAD6-NSmut.
XX      HCV, day non-structural protein; NS3, NS4A, NS4B, NS5A, NS5B;
XX      adenoviral vector; HCV infection; vaccine; gene therapy; protease.
XX      Hepatitis C virus.
XX      Adenovirus serotype 5.
OS      Adenovirus serotype 6.
OS      Synthetic.
XX      Key      Location/Qualifiers
XX      CDS      1264..7221
XX      FT      /tag= 8
XX      FT      /product= "Met-NS3-NS4A-NS4B-NS5A-NS5B (inactive)"
XX      PN      WO2003031588-A2.
XX      PD      17-APR-2003.
XX      PF      10-OCT-2002; 2002WO-US032512.
XX      PR      11-OCT-2001; 2001US-0328655P.
XX      PR      13-MAR-2002; 2002US-0363774P.
XX      PA      (MER) ) MERCK & CO INC.
XX      PA      (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX      PI      Emini EA, Kaslow DC, Bett AJ, Shiver JW, Nicolsia A, Lahm A;
XX      PI      Luzzago A, Cortese R, Colloca S;
XX      PI      WPI; 2003-381708/36.
XX      PT      P-PSDB; AB009574.
XX      PS      New nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A-NS5B polypeptide,
XX      PS      useful as a component of an adenovector or DNA plasmid vaccine for
XX      PS      preventing or treating hepatitis C virus.
XX      PS      Claim 32; Fig 4; 231bp; English.
XX      CC      The invention relates to a nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A
XX      CC      -NS5B (NS stands for non-structural protein from HCV) polypeptide
XX      CC      appearing as AB009574. The encoded polypeptide has sufficient protease
XX      CC      activity to process itself to produce an NS5B protein that is
XX      CC      enzymatically inactive. Also included are a cultured recombinant cell
XX      CC      comprising the novel nucleic acid, making an adenovector (comprising: (a)
XX      CC      producing an adenovirus genome plasmid comprising a gene expression
XX      CC      cassette by homologous recombination between the novel nucleic acid and a
XX      CC      nucleic acid comprising a first adenovirus region from base pair 1-450
XX      CC      corresponding to either Ad5 or Ad6, a second adenovirus region from base
XX      CC      pair 3511-5548 corresponding to Ad5 or from base pair 3508-5541
XX      CC      corresponding to Ad6, joined to the first region, a third adenovirus
XX      CC      region from base pair 5549-28133 corresponding to Ad5 or from base pair
XX      CC      5542-28156 corresponding to Ad6, joined to the second region, a fourth
XX      CC      adenovirus region from base pair 30818-33966 corresponding to Ad5 or from
XX      CC      base pair 30789-33784 corresponding to Ad6, joined to the third region,
XX      CC      and a fifth adenovirus region from base pair 33967-35935 corresponding to
XX      CC      Ad5 or from base pair 33785-35759 corresponding to Ad6, joined to the
XX      CC      fourth region; and (b) rescuing the recombinant adenovirus from the
XX      CC      recombinant adenovirus plasmid), an adenoviral vector that is produced by
XX      CC      method above appearing as ACA61113 which has a humanised version of the
XX      CC      polynucleotide of the invention and encodes the HCV inactivated
XX      CC      polypeptide, a recombinant nucleic acid comprising one or more Ad6
XX      CC      regions and a region not present in Ad6, where at least one Ad6 region is
XX      CC      selected from E1A, E1B, E2A, E2B, E4, E4, E4, E4, E4 and E5. The nucleic acid
XX      CC      is useful as a component of an adenoviral vector or a DNA plasmid vaccine
XX      CC      providing a broad range of antigens for generating an HCV-specific cell
XX      CC      mediated immune response. The nucleic acid may also be used in treating
XX      CC      patients infected with HCV. The present sequence is an adenoviral vector
XX      CC      containing Ad5 and Ad6 sequences together with an HCV promoter and the
XX      CC      non-humanised cDNA encoding the HCV polypeptide with an inactive NS5B

```

CC protein
XX Sequence 37090 BP; 8382 A; 10762 C; 10299 G; 7647 T; 0 U; 0 Other;
SQ Query Match 88.3%; Score 1803; DB 8; Length 37090;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 1 TGGGAGGCGCTTTCACAGGCGCTCAAGCCACGATGAGCCCACTTCTGTCACCAACAAAG 60
DB 2860 TGGGAGAGTGTCTTACAGGCGCTCAAGCCACATAGATGACACTTCTGTGCCAGACAAAG 2919
QY 61 CAGGAGAGAACACTTCCCTTACTGCTGTGGCGTACCAAGCTATGTGTGTGCGTAAAGGCC 120
DB 2920 CAGGAGAGAACACTTCCCTTACTGCTGTGGCGTACCAAGCTATGTGTGTGCGTAAAGGCC 2979
QY 121 CAGGCGCCCACTTCCATCATGGGATCAAAATGTGGAATGTCTCATACGGCTAAAGCGTACT 180
DB 2980 CAGGCGCCCACTTCCATCATGGGATCAAAATGTGGAATGTCTCATACGGCTAAAGCGTACT 3039
QY 181 CTGCGCGGCGCAACACCCCTTGTGTATAGGCTGGAGCGCTCAAAACGAGGTCACCCCTC 240
DB 3040 CTGCGCGGCGCAACACCCCTTGTGTATAGGCTGGAGCGCTCAAAATGAGGTCACCCCTC 3099
QY 241 ACACACCCCATTAACCAATTCATCATGCGATGATGTCAGCCGACCTGAGAGTCTCAAG 300
DB 3100 ACCCAACCCCATTAACCAATTCATCATGCGATGATGTCAGCCGACCTGAGAGTCTCAAG 3159
QY 301 AGCACTTGGGTGCTGTGGCGGGGCTCTTGGAGCTCTGGCGTATGCTTGTGCAACA 360
DB 3160 AGCACTTGGGTGCTGTGGCGGGGCTCTTGGAGCTCTGGCGTATGCTTGTGCAACA 3219
QY 361 GGCAGCGTGTGATTTGGGTAGATTCATCTTGTCCGGCGGCGCGCTATTTGTTCCCGAC 420
DB 3220 GGCAGCGTGTGATTTGGGTAGATTCATCTTGTCCGGAGCGCGCTATTTGTTCCCGAC 3279
QY 421 AGGGAAGTCTCTTACCAAGAGTTCATGATGATGAGAAAGTGGCGGCTGCACTCCCTTAC 480
DB 3280 AGGGAAGTCTCTTACCAAGAGTTCATGATGATGAGAAAGTGGCGGCTGCACTCCCTTAC 3339
QY 481 ATCGAGCAGGGAATGCACTGCCCGAGCGATTCAAGCAAAACCGCTGGGTGCTGCGAC 540
DB 3340 ATCGAGCAGGGAATGCACTGCCCGAGCGATTCAAGCAAAACCGCTGGGTGCTGCGAC 3399
QY 541 ACAGCACAACCAAGCGAGCGCGCTGCTCCGCTGTGAGTTCGAAGTGGCGAGCGCTT 600
DB 3400 ACAGCACAACCAAGCGAGCGCGCTGCTCCGCTGTGAGTTCGAAGTGGCGAGCGCTT 3459
QY 601 GAGACCTTCTGCGGAAACACATGTGGAACCTTCATCAGCGGGAATACATTTAGCAGGC 660
DB 3460 GAGACCTTCTGCGGAAACACATGTGGAACCTTCATCAGCGGGAATACATTTAGCAGGC 3519
QY 661 TTGTTCACCTCTGCTGGGAATCCCGGATTTGCATCATGAGCGTTCAACCTCTGTGC 720
DB 3520 TTATCCACTCTGCTGGGAACCCCGCAATGATCATGATGAGCATTCACACCTCTTATC 3579
QY 721 ACTAAGCCGCTCAACCAACCAATTCCTCTGCTTAAATCATCTGGGAGGATGAGTACGC 780
DB 3580 ACTAAGCCGCTCAACCAACCAATTCCTCTGCTTAAATCATCTGGGAGGATGAGTACGC 3639
QY 781 GCCCAACTGCTGCCCGCAGTGTGCTTACCTTTGTAAGCGCGCGCATTTGCTGTGCG 840
DB 3640 GCCCAACTGCTGCCCGCAGTGTGCTTACCTTTGTAAGCGCGCGCATTTGCTGTGCG 3699
QY 841 GCTGTGTGAGAGATAGGCTTTGGGAAGTGTCTTGTGACATCTTGGCGGGCTATGAGAGA 900
DB 3700 GCTGTGTGAGAGATAGGCTTTGGGAAGTGTCTTGTGACATCTTGGCGGGCTATGAGAGA 3759
QY 901 GGAAGTGGCAGGCGCTGCTGAGCTTTAAGTTCATAGCGCGAATTTGCTCCACCGAG 960
DB 3760 GGAAGTGGCAGGCGCTGCTGAGCTTTAAGTTCATAGCGCGAATTTGCTCCACCGAG 3819
QY 961 GACTGTGTTAACTTACTCTCCCTGCAATCTCTCTCTGTGTCCTGTGCTGGGAGTGTG 1020

DB 3820 GACTGTGTTAACTTACTTCTCCCTGCAATCTCTCTCTGTGTCCTGTGCTGGGATCTGTG 3879
QY 1021 TGGCAGGAGATTAAGTCCGTCGCGACAGTGGGTCCAGGAGAGGGGCTGTGACATGATGAAC 1080
DB 3880 TGTGCAAGAAATACGTGCTGACACGTGGGTCCGGAGAGGGGGCTGTGACATGATGAAC 3939
QY 1081 CGGTGATAGGCTTGCCTGCGGAGGTAAACATGTTTCCCGACGACTATGTGCAAG 1140
DB 3940 CGGTGATAGGCTTGCCTGCGGAGGTAAATCATGTTTCCCGACGACTATGTGCTAG 3999
QY 1141 AGCGACCGCGAGACACGTGTACATTCCTTCCGACCTTATCATACCAACTGTG 1200
DB 4000 AGCGACCGCGAGCGCGTGTACTAGATCTTCCACCTTACATCATGACTGACTG 4059
QY 1201 AAGAGCTCCACAGATGATTAACAGAGACTGTCCACGCGCTGCTCGGCTGAGCTA 1260
DB 4060 AAGAGCTCCACAGATGATTAACAGAGACTGTCCACGCGCTGCTCGGCTGAGCTA 4119
QY 1261 AAGGATGTTTGGGACTGGAATAGCAAGTTTGTGCTGACCTTCAAGACCTGTGCTCAGTCC 1320
DB 4120 AAGGATGTTTGGGACTGGAATAGCAAGTTTGTGCTGACCTTCAAGACCTGTGCTCAGTCC 4179
QY 1321 AAGCTTCTGCGCGATTAACGGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAGGGG 1380
DB 4180 AAGCTTCTGCGCGAGCTACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAGGGG 4239
QY 1381 GTCTGCGGAGAGACGCGATCATGAGACCACTGCTATGTGAGCAGATCACCGGA 1440
DB 4240 GTCTGCGGAGAGACGCGATCATGAGAACACCTGCTATGTGAGCAGATCACCGGA 4299
QY 1441 CATGTCAAAAACGGTTCATAGAGATGTTGCGCTTAAGACCTGTATGATATGTCAT 1500
DB 4300 CATGTCAAAAACGGTTCATAGAGATGTTGCGCTTAAGACCTGTGATGATATGTCAT 4359
QY 1501 GGAATATTCCTCATTAAGCATACACAGCGGCGCTGCAAGCGCTCCCGACGCGCAAC 1560
DB 4360 GGAATATTCCTCATTAAGCATACACAGCGGCGCTGCAAGCGCTCCCGACGCGCAAC 4419
QY 1561 TATTCAGGAGCGCTGTGCGGAGTCTGTAAGAGTACGTGAGATTACCGGAGTGGGG 1620
DB 4420 TATTCAGGAGCGCTGTGCGGAGTCTGTAAGAGTACGTGAGATTACCGGAGTGGGG 4479
QY 1621 GATTTCACTACGTAGCAGCATGACATGACAAAGTAAAGTGCCTGACAGGTTCCA 1680
DB 4480 GATTTCACTACGTAGCAGCATGACATGACAAAGTAAAGTGCCTGACAGGTTCCG 4539
QY 1681 GCCCGCAATTTCTTACAGAAAGTGAATGGGTGCGGCTGCAAGGTTAGCGTCCGGCGTGC 1740
DB 4540 GCTCTGAATTTCTTACAGAAAGTGAATGGGTGCGGCTGCAAGGTTAGCGTCCGGCGTGC 4599
QY 1741 AAACCTCTCTTACGAGGAGAGGTCAATTTCCAGTCCGAGCTCAACCAATACCTGTGGG 1800
DB 4600 AAGCTTCTCTTACGAGGAGAGGTCAATTTCCAGTCCGAGCTCAACCAATACCTGTGGG 4659
QY 1801 TCGAGCTTCCATGCGAGCGCGAACCGAGATGAGAGTGTCACTTTCATAGTCAACGAC 1860
DB 4660 TCGAGCTTCCATGCGAGCGCGAACCGAGATGAGAGTGTCACTTTCATAGTCAACGAC 4719
QY 1861 CCTTCCCATTCACAGCAGACCGCTTAAGCGCAGGCTGCGCAGGGGGTCTCCCGCTCC 1920
DB 4720 CCTTCCCATTCACAGCAGACCGCTTAAGCGCAGGCTGCGCAGGGGGTCTCCCGCTCC 4779
QY 1921 TTGGCAGAGCTTTCAGTACGAGATTTGTCTGGGCTTCTCTCGAAGGCGACATATTAAC 1980
DB 4780 TTGGCAGAGCTTTCAGTACGAGATTTGTCTGGGCTTCTCTCGAAGGCGACATATTAAC 4839
QY 1981 CAATAATGACTTCCAGACGCTGACCTCATGAGGCAACCTTCTGTGGCGGATAGATG 2040
DB 4840 CAATATGACTTCCAGACGCTGACCTCATGAGGCAACCTTCTGTGGCGGAGATAG 4899
QY 2041 GGC 2043
|||

DB 4900 GGC 4902

RESULT 6
ID ADR38452 Standard; DNA; 7987 BP.
XX ADR38452;
XX 18-NOV-2004 (first entry)
XX
XX DNA encoding Hepatitis C virus (isolate BK) NS3-NS5B polypeptide Seq 3.
XX
XX Hepatitis C virus; HCV; NS3; NS5A; non-structural protein;
XX
XX virus replication activity; gene; ds; mutant.
XX
XX Hepatitis C virus (isolate BK).
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX 5'UTR 1..386
XX /tag= a
XX CDS 387..1181
XX /tag= b
XX /product= "Beta-lactamase"
XX misc_feature 1225..1800
XX /tag= c
XX /label= EMCV IRES
XX /note= "internal ribosome entry site"
XX CDS 1801..7758
XX /tag= d
XX /product= "NS3-NS5B polypeptide"
XX /gene= "NS3 (1804-3696)"
XX /gene= "NS4A (3697-3858)"
XX /gene= "NS4B (3859-4641)"
XX /gene= "NS5A (4642-5982)"
XX /gene= "NS5B (5983-7755)"
XX 3'UTR 7759..7987
XX /tag= e
XX
XX WO2004074507-A2.
XX
XX 02-SEP-2004.
XX
XX 09-FEB-2004; 2004WO-US003726.
XX
XX 13-FEB-2003; 2003US-0447318P.
XX
XX (MERI) MERCK & CO INC.
XX
XX Grobler J, Flores O, Markel EJ;
XX
XX MPI; 2004-635590/61.
XX P-PSDB; ADR38450.
XX
XX Making Hepatitis C virus (HCV) replicon having increased replication
XX activity, useful in HCV research, comprises modifying HCV replicon
XX construct to encode an amino acid substitution at a position
XX corresponding to amino acid 470 of NS3.
XX
XX Claim 10; SEQ ID NO 3; 54pp; English.
XX
XX This invention relates to a novel method for producing a Hepatitis C
XX virus (HCV) replicon having an increased replication activity.
XX Specifically, it refers to modifying an HCV replicon construct to encode
XX an amino acid substitution in NS3 (a non-structural protein that along
XX with NS4A, NS4B, NS5A and NS5B make up the virus replication machinery
XX released in the form of a polypeptide). The present invention describes
XX an amino acid substitution at a position corresponding to amino acid 470
XX of NS3 alone, or in combination with, an isoleucine in a position
XX corresponding to amino acid 232 of NS5A that confers improved cell
XX culture replication activity compared to wild type HCV. The method is
XX useful for facilitating the identification of broadly efficacious

CC compounds against different HCV isolates and facilitating HCV research,
CC where compounds that inhibit HCV replication have research and
CC therapeutic applications in identifying overall efficacy and lack of
CC unacceptable toxicity. Accordingly, they can be used to treat or inhibit
CC the onset of HCV in a patient. This polynucleotide is a mutant HCV
CC replicon DNA sequence that encodes an HCV NS3-NS5B polypeptide (based on
CC HCV-BK) with 471M and 1179I residue substitutions, given in an
CC exemplification of the invention.
XX
XX Sequence 7987 BP; 1722 A; 2315 C; 2204 G; 1746 T; 0 U; 0 Other;
XX
XX Query Match 88.2%; Score 1801.4; DB 13; Length 7987;
XX Best Local Similarity 92.6%; Pred. No. 0;
XX Matches 1892; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
XX
QY 1 TGGGAGGGGGCTTTCACAGGCGCTCCACCGGATGAGGCGGCTTCCGCAACAAAG 60
DB 3397 TGGGAGAGTGTCTTACAGGCGCTCCACCGATAGAGACACTTCTTCCAGACAAAG 3456
QY 61 CAGGACGAGAGCAACTTCCCTACCTGTGGCGTACAGGCTATGCTGCGCTAGGGCC 120
DB 3457 CAGGACGAGAGCAACTTCCCTACCTGTGGCGTATGAGATACAGGCGTGGCGGCGCT 3516
QY 121 CAGGCGCCACCTTCATCATGAGATTAATGTGAAAGTGTCTATACGCTTAAAGCTTACT 180
DB 3517 CAGGCGCCACCTTCATCATGAGATTAATGTGAAAGTGTCTATACGCTTAAAGCTTACG 3576
QY 181 CTGGCGCGGCGCAACACCTTGTGTATAGGCTGGGAGCGGCTCAAAAGAGGTCACCCCTC 240
DB 3577 CTGCAAGGGGCAACACCTTGTGTATAGGCTGGGAGCGGCTCAAAAGAGGTCACCCCTC 3636
QY 241 ACACACCCCATTAACCAATTTCATCATGAGATGATGATGAGGCTGAGGCTGAGGCTGCTACG 300
DB 3637 ACCCAACCCCATTAACCAATTTCATCATGAGATGATGATGAGGCTGAGGCTGAGGCTGCTAC 3696
QY 301 AGCACCTGGGCTCTGTGTGGCGGGGCTCTTGAGGCTTGGCTGCTGATTTGCTTGACACACA 360
DB 3697 AGCACCTGGGCTCTGTGTGGCGGGGCTCTTGAGGCTTGGCTGCTGATTTGCTTGACACACA 3756
QY 361 GGCAGCGTGTGATTTGTGGTATGATCATCTTGTCCGGCGCGGCTATTTGTTCCCGAC 420
DB 3757 GGCAGCGTGTGATTTGTGGTATGATCATCTTGTCCGGCGCGGCTATTTGTTCCCGAC 3816
QY 421 AGGGAAGTCTCTTACACAGAGTTGATGATGAGAAAGTGGCGGCTGACCTCTCTTAC 480
DB 3817 AGGGAAGTCTCTTACACAGAGTTGATGATGAGAAAGTGGCGGCTGACCTCTCTTAC 3876
QY 481 ATGAGCAGGAGATGAGCTTCCCGAGCAGTTCAAGCAAAAAGCGCTCGGTTGCTGACG 540
DB 3877 ATGAGCAGGAGATGAGCTTCCCGAGCAGTTCAAGCAAAAAGCGCTCGGTTGCTGACG 3936
QY 541 ACAGCCACCAAGCAAGCGAGCGGCTGCTCCCGTGTGTGAGTCCAAATGGGAGCGCTT 600
DB 3937 ACAGCCACCAAGCAAGCGAGCGGCTGCTCCCGTGTGTGAGTCCAAATGGGAGCGCTT 3996
QY 601 GAGACCTTGGGGGCAACATGAGAACTTCATCAGGAGGATACAGTACTTAGCAGGC 660
DB 3997 GAGACCTTGGGGGCAACATGAGAACTTCATCAGGAGGATACAGTACTTAGCAGGC 4056
QY 661 TTGTCACTCTGCTGGGATCCCGGATGATCATCATGATGAGGCTTCAAGCCTCTGTC 720
DB 4057 TTGTCACTCTGCTGGGATCCCGGATGATCATCATGATGAGGCTTCAAGCCTCTGTC 4116
QY 721 ACTAGCCGCTCAACCAATTCATCCTTCTGCTTAAATCTCTGGGGGATGGGTAGGC 780
DB 4117 ACTAGCCGCTCAACCAATTCATCCTTCTGCTTAAATCTCTGGGGGATGGGTAGGC 4176
QY 781 GCCCAACTGCTCCCGGAGTGTGCTGAGCTTGTGAGGCGCGGCTTCTGCTGCTG 840
DB 4177 GCCCAACTGCTCCCGGAGTGTGCTGAGCTTGTGAGGCGCGGCTTCTGCTGCTG 4236
QY 841 GCTGTGGAGAGTATGAGGCTTGGGAAAGTCTTGTGACATCTTGGCGGCTATGAGAGA 900

Db	4237	GCTGTGGACAGTATAGGCTTGGGAAGTGCTGTGTGACATTCTGGCGGGTTATGAGCA	4238
QY	901	GGAGTGGCAGGCGCGCTCGTGGCCCTTTAAGGTCATGACGGCGGAATGCTCCACCGAG	960
Db	4297	GGAGTGGCCGGGCGCTGCTGGCCCTTCAAGGTCAAGACGGCGAGATGCTCCACCGAG	4356
QY	961	GACCTGGTTAACTTACTCTCCCTGGCATCTCTCTCTCGGTGGCCCTGGTGGTGGGTGGTG	1020
Db	4357	GACCTGGTCAATCTACTCTTCGCAATCTCTCTCTGGCGCCCTGGTGTGGGGGTCTGTG	4416
QY	1021	TGCGAGAGATACTCTCGTCGGGACCTGGAGTCCAGGGGAGGGGGCTGTSCATGTGATGAC	1080
Db	4417	TGTGAGACAAATCTGCTGTGACACTGGGGTCCGGAGAGGGGGCTGTGTGATGTATGAC	4476
QY	1081	CGGCTGATAGCGTTGGCTTCGGGGGGTAAACCATGTTTCCCCACGCACTATGTGCAAG	1140
Db	4477	CGGCTGATAGCGTTGGCTTCGGGGGGTAAATCATGTTTCCCCACGCACTATGTGCTGAG	4536
QY	1141	AGCGACGGCGGACGACCGTGTCACTAGATCTCTTCCGACCTTATCTATACCCAACTGTG	1200
Db	4537	AGCGACGGCGGACCGCGTGTACTAGATCTCTTCCACCTTAACTATCACTCACTGTGTG	4596
QY	1201	AAGAGGCTCCACCACTGGATTTAAGAGGACTGCTCAGCGCTGTCCGGTCCGGGTGA	1260
Db	4537	AAAGAGCTCCACCACTGGATTTATATBAAGACTGTCTCAACCGTGTTCGGGCTCGTGCTA	4656
QY	1261	AGGATGTTTGGGACTGTGATATGCAAGTTTTGGCTGTACTTCAAGACCTGTGCTCAGTCC	1320
Db	4657	AGGATGTTTGGGACTGTGATATGACCGGTGTGTACTGTACTTCAAGACCTGTGCTCAGTCC	4716
QY	1321	AAGCTCTGCGCGCATTAACCGGAGTCCCTTTTCTATGTCAACTGTGGTCAAGGGG	1380
Db	4717	AAGCTCTGCGCGCACTACCGGAGTCCCTTTTCTGTGTCAACCGGGGTAAAGAGGA	4776
QY	1381	GTCTGGGGGGGAGAGGGATCATGAGACACTGTCTATGTGGAGCAATATCCGGGA	1440
Db	4777	GTCTGGGGGGGAGAGGGATCATGTGAACACTGTGCCATGTGGAGCAATATCCGGGA	4836
QY	1441	CATGTCAAAAAAGTTCATGATGAGATCGTTGGGCTTAAGACCTGTATGATGATGTGCAT	1500
Db	4837	CATGTCAAAAAAGTTCATGATGAGATCGTGGGCTTAAGACCTGTGAGCAACGTGGCAT	4896
QY	1501	GGAACATTCCCATCATACGCATACACACGGGCCCCCTGACGCTTCCCAAGCGCAAC	1560
Db	4897	GGAACATTCCCATCATACGCATACACACGGGCCCCCTGACACCTTCCAGCGCAAC	4956
QY	1561	TATTCAGGGGCGCTGTGGCGGGTGGCTGTAGAGATACGTGGAGTTACCGGGTGGGG	1620
Db	4957	TATTCAGGGGCGCTGTGGGGGGTGGCGCTGTAGAGATACGTGGAGTTACCGGGTGGGG	5016
QY	1621	GATTTTCACTAAGTACGACGATGACCACTGACAACTGTAATAATGCCCTGTCCAGTTTCCA	1680
Db	5017	GATTTTCACTAAGTACGACGCGGATGACCACTGACAACTGTAATAATGTCAGAGTTCCG	5076
QY	1681	GCCCCGAATTTCTTACAGAAGTGATGGGGTGGCTGCAAGGTACGTTCCGGGCTGC	1740
Db	5077	GCTCTGAATTTCTTACCGGAGGTGACCGGAGTGGCGGTTGCAACAGTACGTTCCGGGCTGC	5136
QY	1741	AAACCTCTCTAAGGAGAGAGGTTCACATTCCAGGTCCGGGCTCAACCAATACCTGTTGGG	1800
Db	5137	AGGCTCTCTCTAAGGAGAGAGGTTCACATTCCAGGTCCGGGCTCAACCAATACCTGTTGGG	5196
QY	1801	TGCGAGCTCCCATGTGGAGGCCGAACCGGATGTAGAGAGTCTCACTTCATGCTCACGAC	1860
Db	5197	TGCGAGCTTACATGTGAGGCCGAACCGGATGTAGAGAGTCTCACTTCATGCTCACGAC	5256
QY	1861	CCCTTCCCATATCAGACAGACAGCGCTAAGCGCAGGCTGGCCAGGGGGTCTTCCCTTCC	1920
Db	5257	CCCTTCCCATATCAGACAGAAACGGCTAAGCGTAGGTTGGCCAGGGGGTCTTCCCTTCC	5316
QY	1921	TTGGCAGAGTCTTTCAGCTTACGCAAGTTGTCTGGGCTTCTCGAAGGCGACATCACTTACC	1980
Db	5317	TTGGCAGAGTCTTTCAGCTTATCAGTGTGTCTGGGCTTCTTGAAGGCGACATGCACTTACC	5376

QY	1961		AAAAATGACTTCCGACAGCGCTGACTTATCGACGCCAAACCTTCGTGGCGGCATGAGATG	204
Db	5377	CACCATGCTCTCTCGAGCGCTGACCTCATCGAGCCAACTTCGTGTGGCGGAGAGATG	5436	
QY	2041	GGC	2043	
Db	5437	GGC	5439	
RESULT 7				
ADRs2189				
ID	ADRs2189	standard; DNA; 9587 BP.		
AC	ADRs2189;			
XX				
DT	16-DEC-2004	(first entry)		
XX				
DE	Hepatitis C virus type 1b polyprotein DNA.			
XX				
KM	antihypemic; cardiact; vaeotropic; antiarteriosclerotic; antidiabetic;			
KM	cytostetic; anticonvulsant; nootropic; muscular; anti-HIV;			
KM	RNA interference; iRNA; antisense technology; lipid metabolism;			
KM	cholesterol imbalance; dyslipidaemia hypercholesterolaemia;			
KM	coronary artery disease; CAD; coronary heart disease; CHD;			
KM	atherosclerosis; hepatic glucose production;			
KM	glucose-metabolism-related disorder; diabetes; cancer; breast cancer;			
KM	colon cancer; lung cancer; neurological disease; Huntington disease;			
KM	sphinocebellar ataxia; viral disease; AIDS; HCV; polyprotein; gene; ds.			
OS	Hepatitis C virus.			
XX				
XX	WO2004080406-A2.			
PN				
XX	23-SEP-2004.			
PD				
XX				
PF	08-MAR-2004; 2004WO-US007070.			
XX				
XX	07-MAR-2003; 2003US-0452682P.			
PR	12-MAR-2003; 2003US-0454265P.			
PR	13-MAR-2003; 2003US-0454962P.			
PR	13-MAR-2003; 2003US-0455050P.			
PR	14-APR-2003; 2003US-0462894P.			
PR	17-APR-2003; 2003US-0463772P.			
PR	25-APR-2003; 2003US-0465665P.			
PR	25-APR-2003; 2003US-0465802P.			
PR	09-MAY-2003; 2003US-0469612P.			
PR	08-AUG-2003; 2003US-0493986P.			
PR	11-AUG-2003; 2003US-0494597P.			
PR	26-SEP-2003; 2003US-0506341P.			
PR	09-OCT-2003; 2003US-0510246P.			
PR	10-OCT-2003; 2003US-0510318P.			
PR	07-NOV-2003; 2003US-0518453P.			
XX				
PA	(ALNY-) ALNYLAM PHARM.			
PI				
PI	Manoharan M, Bumcrot D;			
XX				
XX	WPI; 2004-677362/66.			
XX				
PT	Interference RNA agent useful for treating dyslipidaemias, coronary artery			
PT	diseases, diabetes, cancer or neurological disease, comprises sense			
PT	sequence and antisense sequence which has specific modifications.			
XX				
XX	Example 5; SEQ ID NO 6688; 378pp; English.			
PS				
XX				
CC	The invention describes a RNA interference (iRNA) agent (I) comprising a			
CC	sense sequence and an antisense sequence, where the sense sequences have			
CC	one or more asymmetrical 2'-O alkyl modifications, the antisense			
CC	sequences have one or more asymmetrical phosphorochiaste modifications			
CC	and the antisense sequence targets a human gene sequence. Also described			
CC	are a pharmaceutical preparation comprising (I); reducing (MI) apob-100			
CC	levels or glucose-6-phosphataae levels in a subject; producing (I);			

CC stabilising (1), involves selecting a sequence with activity and
 CC introducing one or more asymmetrical modification in the sequence, where
 CC the modification decreases nuclease sensitivity while not decreasing its
 CC activity; a kit comprising (I) and instruction for its use; and a device
 CC that can be dispense or administer a composition comprising (I). (I) is
 CC useful for reducing apob-100 levels or glucose-6-phosphatase levels. (M1)
 CC is useful for reducing apob-100 levels or glucose-6-phosphatase levels.
 CC The subject is suffering from a disorder characterised by elevated or
 CC otherwise unwanted expression of apob-100, elevated or otherwise unwanted
 CC levels of cholesterol, and/or dysregulation of lipid metabolism. The
 CC disorder is chosen from the HDL/LDL cholesterol imbalance,
 CC hyperlipidaemia, hypercholesterolaemia, statin-resistant
 CC dyslipidaemia, coronary artery disease (CAD), coronary heart
 CC disease (CHD) and atherosclerosis. (I) is administered to a subject to
 CC inhibit hepatic glucose production or for treating glucose-metabolism-
 CC related disorder e.g. diabetes or type-2 diabetes. (I) is useful for
 CC treating the diseases as mentioned above, cancer (e.g. breast, colon or
 CC lung cancer), neurological disease (e.g., Huntington disease or
 CC spinocerebellar ataxia) or viral disease (e.g., AIDS). This sequence
 CC represents hepatitis C virus type 1b polypeptide DNA.

XX Sequence 9587 BP; 1921 A; 2865 C; 2714 G; 2087 T; 0 U; 0 Other;

Query Match 88.2%; Score 1801.4; DB 13; Length 9587;
 Best Local Similarity 92.6%; Pred. No. 0;
 Matches 1892; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 TGGGAGGCGCTTTCACAGGCTCCACCCAGTGGATGCCACTTCTGTCACCAACAAAG 60
 Db 5013 TGGGAGAGTCTTTCACAGGCTCCACCCAGTGGATGCCACTTCTGTCACCAACAAAG 5072
 QY 61 CAGGAGGAGACAACTTCCCTACCTGATGAGGCTACAGGCTACTGTCGCTAAGGCGC 120
 Db 5073 CAGGAGGAGACAACTTCCCTACCTGATGAGGCTACAGGCTACTGTCGCTAAGGCGC 5132
 QY 121 CAGGCGCCCACTTCATCATGAGATCAATGTGAAAGTGTCTCATACGCTTAAAGCTTACT 180
 Db 5133 CAGGCGCCCACTTCATCATGAGATCAATGTGAAAGTGTCTCATACGCTTAAAGCTTACT 5192
 QY 181 CTGGCGGGGCGCAACCTTCTGCTGATGAGGCTGGAGCCGCTCCAAACAGAGTCACTTC 240
 Db 5193 CTGGCGGGGCGCAACCTTCTGCTGATGAGGCTGGAGCCGCTCCAAACAGAGTCACTTC 5252
 QY 241 ACACACCCCATTAACAAATTTCATGAGGATGATGATGATGATGATGATGATGATGATGAT 300
 Db 5253 ACCCACCCCATTAACAAATTTCATGAGGATGATGATGATGATGATGATGATGATGATGAT 5312
 QY 301 AGCACCTGGGTGCTGATGAGGCGGGGCTCTTGACAGCTTGGCTGCTGATTTGTCACACA 360
 Db 5313 AGCACCTGGGTGCTGATGAGGCGGGGCTCTTGACAGCTTGGCTGCTGATTTGTCACACA 5372
 QY 361 GGCACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 Db 5373 GGCACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5432
 QY 421 AGGGAAGTCTCTACAGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 Db 5433 AGGGAAGTCTCTACAGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 5492
 QY 481 ATCGAGAGGGAATCAGCTCCGAGAGAGTTCAGAGCAAAAGAGCTCGGGTTGTCAGAG 540
 Db 5493 ATCGAGAGGGAATCAGCTCCGAGAGAGTTCAGAGCAAAAGAGCTCGGGTTGTCAGAG 5552
 QY 541 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCGCTGATGAGTCCAGTGGCGAGGCTT 600
 Db 5553 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCGCTGATGAGTCCAGTGGCGAGGCTT 5612
 QY 601 GAGACCTTCTGGGCGAAACACATGTGAACTTCATCAGCGGAGTACAGTACTTGAAGAGC 660
 Db 5613 GAGACCTTCTGGGCGAAACACATGTGAACTTCATCAGCGGAGTACAGTACTTGAAGAGC 5672
 QY 661 TTGTCACTCTGCTGAGGATCCGCGATGTGATCATGATGAGCTTCAAGCCTCTGTCT 720

Db 5673 TTATCCACTCTGCTGGGAAACCCCGCAATGATCATGATGATGATGATGATGATGATGATGAT 5732
 QY 721 ACTAGCCGCTACACCAACCAATTCACCTCCGTTAATCATCTGAGGGGATGGGATGCC 780
 Db 5733 ACTAGCCGCTACACCAACCAATTCACCTCCGTTAATCATCTGAGGGGATGGGATGCC 5792
 QY 781 GCCCACTGCTCCCGCAGAGTCTCTCAGCTTTCATGAGGCGCGGCAATTCGATGTCG 840
 Db 5793 GCCCACTGCTCCCGCAGAGTCTCTCAGCTTTCATGAGGCGCGGCAATTCGATGTCG 5852
 QY 841 GCTGTGGCAGATAGGCTTGGGAAAGTCTTTCATGAGTTCATGAGTTCATGAGTTCATGAG 900
 Db 5853 GCTGTGGCAGATAGGCTTGGGAAAGTCTTTCATGAGTTCATGAGTTCATGAGTTCATGAG 5912
 QY 901 GGAAGTGGAGGCGGCTGCTGAGCTTTCATGAGTTCATGAGTTCATGAGTTCATGAGTTCATGAG 960
 Db 5913 GGAAGTGGAGGCGGCTGCTGAGCTTTCATGAGTTCATGAGTTCATGAGTTCATGAGTTCATGAG 5972
 QY 961 GACCTGTTCATTCATCTCCCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 5973 GACCTGTTCATTCATCTCCCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6032
 QY 1021 TGGCAGCGATATCTGCTGCGACGCTGCGTCCAGGAGGAGGCTGTCAGTGTGATGAC 1080
 Db 6033 TGGCAGCGATATCTGCTGCGACGCTGCGTCCAGGAGGAGGCTGTCAGTGTGATGAC 6092
 QY 1081 CGGCTGATAGCTTTCGCTGCGGCGGCTGACATGTTTCCCAACCACTATGTCAGAG 1140
 Db 6093 CGGCTGATAGCTTTCGCTGCGGCGGCTGACATGTTTCCCAACCACTATGTCAGAG 6152
 QY 1141 AGCGAGCGCGAGCAGTGTCTCAGATTCCTCTCCGACTTATCATCAACCACTGTTG 1200
 Db 6153 AGCGAGCGCGAGCAGTGTCTCAGATTCCTCTCCGACTTATCATCAACCACTGTTG 6212
 QY 1201 AAGAGCTTCACCACTGATTTAACAGAGACCTGCTCACGCCCTGCTCGGCTGTGCTA 1260
 Db 6213 AAGAGCTTCACCACTGATTTAACAGAGACCTGCTCACGCCCTGCTCGGCTGTGCTA 6272
 QY 1261 AAGATGTTTGGAGTGTGATATGACACGTTTGGTGACTTCAACCTGCTTCCAGTCC 1320
 Db 6273 AAGATGTTTGGAGTGTGATATGACACGTTTGGTGACTTCAACCTGCTTCCAGTCC 6332
 QY 1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCAACGATGATGATGATGATGATGAT 1380
 Db 6333 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCATGCAACGATGATGATGATGATGATGAT 6392
 QY 1381 GTTGGCGGGAGACGCGATCATGACACACTGCTCATGTGAGGACACATCACCGGA 1440
 Db 6393 GTTGGCGGGAGACGCGATCATGACACACTGCTCATGTGAGGACACATCACCGGA 6452
 QY 1441 CATGTCAAAAACGGTTCATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 Db 6453 CATGTCAAAAACGGTTCATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 6512
 QY 1501 GGAACATTCCTCATCAAGCATACACACAGGAGCCCTGCAAGCCTCCACAGGCGCAAC 1560
 Db 6513 GGAACATTCCTCATCAAGCATACACACAGGAGCCCTGCAAGCCTCCACAGGCGCAAC 6572
 QY 1561 TATTCAGAGGCGCTGAGCGGCTGCTGAGAGTACGTGAGGTTACCGGGGTGGGG 1620
 Db 6573 TATTCAGAGGCGCTGAGCGGCTGCTGAGAGTACGTGAGGTTACCGGGGTGGGG 6632
 QY 1621 GATTTTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Db 6633 GATTTTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6692
 QY 1681 GCGCCGGAATCTTCAAGAGTGGATGAGGAGGCGCTGCAAGGATGAGCTCCGCGCTGC 1740
 Db 6693 GCGCCGGAATCTTCAAGAGTGGATGAGGAGGCGCTGCAAGGATGAGCTCCGCGCTGC 6752
 QY 1741 AAACCTCTCTACGAGAGAGTTCATTCAGAGTGGGCTCAACCAATTCATGATGATGATGATGAT 1800
 Db 6753 AAACCTCTCTACGAGAGAGTTCATTCAGAGTGGGCTCAACCAATTCATGATGATGATGATGAT 6812


```

Db      5424 AGGGAGCTTCTCTACAGAGATTGCAATGAAAGTAGCGCTCGACCTCCCTTAC 5483
Qy      481 ATCGAGCGGGGAATGAGTCCGCCAGAGCTTCAAGCAAAAAGCGCTCGGTTGCTGAG 540
Db      5484 ATCGAGCGGGGAATGAGTCCGCCAGAGCTTCAAGCAAAAAGCGCTCGGTTGCTGAG 5543
Qy      541 ACAGCACCAGCAAGCGGAGCGCGTGTCTCCGTTGAGTGCAGTGCAGAGCTT 600
Db      5544 ACAGCACCAGCAAGCGGAGCGCGTGTCTCCGTTGAGTGCAGTGCAGAGCTT 5603
Qy      601 GAGACTTCTGAGGCGAAACATGTTGAACTTTCATCAGCGGAGTACAGTACTTACAGGC 660
Db      5604 GAGACTTCTGAGGCGAAACATGTTGAACTTTCATCAGCGGAGTACAGTACTTACAGGC 5663
Qy      661 TTGTTCACCTCTGCTGAGGAAATCCCGGATTTGATCTCATGAGCGCTTACAGCTTGTG 720
Db      5664 TTATTCACCTCTGCTGAGGAAATCCCGGATTTGATCTCATGAGCGCTTACAGCTTGTG 5723
Qy      721 ACTAGCCGCTCACCACCCCAATCTACCTCTGCTTAACTCATCTGAGGAGATGAGTACGC 780
Db      5724 ACCAGCCGCTCACCACCCCAATCTACCTCTGCTTAACTCATCTGAGGAGATGAGTACGC 5783
Qy      781 GCCCACTGCTCCCCCAGTGTCTGTTGAGGCGCGGCAATTGCTGTGCTG 840
Db      5784 GCCCACTGCTCCCCCAGTGTCTGTTGAGGCGCGGCAATTGCTGTGCTG 5843
Qy      841 GCTGTGCGAGCATAGGCTTGGGAAAGTGTGTTGAGATCTTGGCGGCTATGAGAGA 900
Db      5844 GCTGTGCGAGCATAGGCTTGGGAAAGTGTGTTGAGATCTTGGCGGCTATGAGAGA 5903
Qy      901 GGAAGTGGAGGCGCGCTGTGGCTTTAAGTCAATGAGCGGAGAAATGCCCTCCACCGAG 960
Db      5904 GGAAGTGGAGGCGCGCTGTGGCTTTAAGTCAATGAGCGGAGAAATGCCCTCCACCGAG 5963
Qy      961 GACCTGTGTTAACTTACTCCCTGCAATCTCTCTGCTGAGTCCCTGTGCTGCGGCTGTG 1020
Db      5964 GACCTGTGTTAACTTACTCCCTGCAATCTCTCTGCTGAGTCCCTGTGCTGCGGCTGTG 6023
Qy      1021 TCGCGAGCATTACTGCTGCGGAGCTGTGGTCCAGGGAGGGGCTGTGCAATGATGAAC 1080
Db      6024 TGTGAGCAATTACTGCTGCGGAGCTGTGGTCCAGGGAGGGGCTGTGCAATGATGAAC 6083
Qy      1081 CGGCTGATAGCGTTGGCTCGGAGGGGTAACAATTTCCCGACGACTATGACAGAG 1140
Db      6084 CGGCTGATAGCGTTGGCTCGGAGGGGTAACAATTTCCCGACGACTATGAGCTGAG 6143
Qy      1141 AGCGAGCGCGAGCACTGTCTCACTGAGATCTCTCCGACTTACTATCAACCAACTGTG 1200
Db      6144 AGCGAGCGCGAGCACTGTCTCACTGAGATCTCTCCGACTTACTATCACTGAGCTGTG 6203
Qy      1201 AAGAGGCTTCACAGTGAATTAACAAGAGCTGCTCCAGCCCTGCTCGGCTGTGAGTGA 1260
Db      6204 AAGAGGCTTCACAGTGAATTAACAAGAGCTGCTCCAGCCCTGCTCGGCTGTGAGTGA 6263
Qy      1261 AGGAGTGTGGGAGCTGATATGCAAGTGTGGCTGATTAAGAAGCTGCTCCAGTGC 1320
Db      6264 AGGAGTGTGGGAGCTGATATGCAAGTGTGGCTGATTAAGAAGCTGCTCCAGTGC 6323
Qy      1321 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTCAATGCAAGCTGTGATACAAGGG 1380
Db      6324 AAGCTCTGCGCGATTAACCGGAGTCCCTTTTCTGCTGCAAGCGGGATACAAGGGA 6383
Qy      1381 GTCTGGCGGGAGAGCGGATCATGCAACCACTGCTCATGTGAGACACAGATCACCGGA 1440
Db      6384 GTCTGGCGGGAGAGCGGATCATGCAACCACTGCTCATGTGAGACACAGATCACCGGA 6443
Qy      1441 CATGTCAAAAAGGTTTCATGAGGATGTTGGGCTTAAGACCTGTATTAACATGAGGAT 1500
Db      6444 CATGTCAAAAAGGTTTCATGAGGATGTTGGGCTTAAGACCTGTATTAACATGAGGAT 6503
Qy      1501 GGAACTTCCCATCAAGCATACACAGGGGCTCTGACGCGCTCCCAAGGCCCAAC 1560

```

```

Db      6504 GGAACTTCCCATCAAGCATACACAGGGGCTCTGACCACTCTTCCAGGCCCAAC 6563
Qy      1561 TATTCAGGAGCGCTGTGGCGGTTGCTGTGAGAGTACGTGAGGTTACGGCGGTTGAG 1620
Db      6564 TATTCAGGAGCGCTGTGGCGGTTGCTGTGAGAGTACGTGAGGTTACGGCGGTTGAG 6623
Qy      1621 GATTTCACTACGTACAGAGCATGACCACTGCAACGTTAAATGCCCCGTGCCAGTTTCA 1680
Db      6624 GATTTCACTACGTACAGAGCATGACCACTGCAACGTTAAATGCCCCGTGCCAGTTTCCG 6683
Qy      1681 GCCCGCAATTTTTCACAGAGTGAATGAGGCTGCGCTGTGACAGTACGCTCCGCGTGC 1740
Db      6684 GCTCCTGAATTTCTTCGAGAGTGAACGAGTGCAGGCTTGCACAGGATACGCTCCGCTGC 6743
Qy      1741 AAACCTCTCTACGAGAGAGTGCATTTCCAGGTGCGGCTCAACCAATACCTGTGGG 1800
Db      6744 AGGCTCTCTACGAGAGAGTGCATTTCTCAGGTGCGGCTCAACCAATACCTGTGGG 6803
Qy      1801 TCGAGCTTCCATGCGAGAGCCGAAACCGGATGAGAGTGTCACTTCCATGCTCACAGAC 1860
Db      6804 TCAGAGCTACATGAGAGAGCCGAAACCGGATGAGAGTGTCACTTCCATGCTCACAGAC 6863
Qy      1861 CCTTCCCATACAGAGAGAGCGCTTAAAGCGGCTGAGCGGCTGCTCCCTCC 1920
Db      6864 CCTTCCCATACAGAGAGAGCGCTTAAAGCGGCTGAGCGGCTGCTCCCTCC 6923
Qy      1921 TTGGCAGCTCTTACGATGACCATGTTGCTGCGCTCTCGAAGGCGCATTAATTAAC 1980
Db      6924 TTGGCAGCTCTTACGATGACCATGTTGCTGCGCTCTCGAAGGCGCATTAATTAAC 6983
Qy      1981 CAAATGACTTCCAGAGCGCTGACCTCATTCAGAGCCAACTTCTGTGCGGATGAGT 2040
Db      6984 CACCATGCTCTCCGAGAGCTTACCTCATTCAGAGCCAACTTCTGTGCGGATGAGT 7043
Qy      2041 GGC 2043
Db      7044 GGC 7046

RESULT 9
ID      AAQ21829
XX      AAQ21829 standard; cDNA; 9416 BP.
AC      AAQ21829;
XX      XX
DT      25-MAR-2003 (revised)
DT      01-MAY-1992 (first entry)
XX      XX
DE      Non-A, non-B viral genome.
XX      XX
KW      NANBV; vaccine; immunodiagnosis; antigen; antibody; ds.
XX      XX
OS      Non-A.
OS      Non-B hepatitis virus.
XX      XX
FH      Key
FT      CDS      location/Qualifiers
FT      FT      333..677
FT      FT      /*tag= a
FT      FT      /product= "C"
FT      FT      /note= "core protein"
FT      FT      678..905
FT      FT      /*tag= b
FT      FT      /product= "W"
FT      FT      /note= "matrix protein"
FT      FT      906..1499
FT      FT      /*tag= c
FT      FT      /product= "E"
FT      FT      /note= "envelope protein"
FT      FT      1500..2519
FT      FT      /*tag= d
FT      FT      /product= "NS1"
FT      FT      2520..3350
FT      FT      /*tag= e
FT      CDS

```



```

FT FT /product= "NS2"
FT CDS 3351..5177
FT /tag= f
FT /product= "NS3"
FT CDS 5178..5918
FT /tag= g
FT /product= "NS4a"
FT CDS 5919..6371
FT /tag= h
FT /product= "NS4b"
FT CDS 6372..9365
FT /tag= i
FT /product= "NS5"
XX EP464287-A.
XX PN
XX PD
XX 08-JAN-1992.
XX 28-DEC-1990; 90BP-00314371.
XX PF
XX 25-JUN-1990; 90JP-00167466.
XX PR 31-AUG-1990; 90JP-00230921.
XX PR 09-NOV-1990; 90JP-00305605.
XX PR 17-JUN-1991; 91BP-00401604.
XX PA
XX (OSAU ) UNITV OSAKA.
XX DR WPI, 1992-009617/02.
XX P-PSDB: AAR20091.
XX PT New DNA from non-A, non-B hepatitis virus - and derived antigenic
XX polyptide(s) useful for diagnostics, blood screening and in vaccines.
XX
XX PS Claim 1; Fig 2; 89pp; English.
XX
CC The sequence was obtd. from several overlapping "BK" cDNA clones obtd. by
CC "gene walking" using a cDNA clone isolated from a library prepd. from
CC NANBV RNA. The DNA and fragments of it can be used for the detection of
CC the presence of NANBV by hybridisation or PCR. Antigenic polyptides
CC encoded by the sequence can be used as immunoassay reagents, for
CC screening donated blood, and as immuno- genes for vaccine prodn.
CC Antibodies raised to the peptides can be used in immunoassays to detect
CC or quantify NANBV antigens in liver tissue and blood. Preferred
CC polyptides are encoded by the following nucleotides: 333-422, -677, or
CC -631; 474-563; 678-905; 906-953 or -1499; 1020-1046 or -1121; 1194-1232;
CC 1209-1322; 1500- 2519; 2520-3350; 3351-5177; 4485-4574; 5178-5918; 5544-
CC 5633; 5919- 6371; and 6372-9372. The sequence is also disclosed in BP-
CC 463848 in which a virus particle cong. antigens encoded by the sequence
CC is claimed, as well as expression vectors contg. the sequence. See
CC AAQ20268 for details of this specification. (Updated on 25-MAR-2003 to
CC correct PA field.)
XX
SQ Sequence 9416 BP; 1905 A; 2825 C; 2679 G; 2007 T; 0 U; 0 Other;
Query Match 87.9%; Score 1796.6; DB 2; Length 9416;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```

```

QY 241 AACAACCCATTAACCAATTATCATGAGCATGATGATGACCGACCTGGAGTGTACG 300
DB 5244 ACCACCCCATTAACCAATTATCATGAGCATGATGATGACCGACCTGGAGTGTACG 5303
QY AGCACTGGGTGCTGGTGGGGGGGCTCTTACAGCTGCTGGCTATTTGCTTGACAACA 360
DB 5304 AGCACTGGGTGCTGGTGGGGGGGCTCTTACAGCTGCTGGCTATTTGCTTGACAACA 5363
QY 361 GGCAGCGTGTATTTGTTGGTGGATCATCTTGTCCGGGCGGGCTATTTGTTCCGAC 420
DB 5424 AGGGAGCTTCTTACAGAGAGTTGATGATGAAAGTGGCGGTGCTCCACTCCCTTAC 480
QY 421 AGGGAATCTTCTTACAGAGAGTTGATGATGAAAGTGGCGGTGCTCCACTCCCTTAC 480
DB 5424 AGGGAGCTTCTTACAGAGAGTTGATGATGAAAGTGGCGGTGCTCCACTCCCTTAC 5483
QY 481 ATCGAGCAGGGAATGAGCTGCGGAGCATGTTCAAGCAAAAGCGCTGGGTGCTGAG 540
DB 5484 ATCGAGCAGGGAATGAGCTGCGGAGCATGTTCAAGCAAAAGCGCTGGGTGCTGAG 5543
QY 541 ACAGCCACCAAGCAAGCGAGCGCTGCTCCGTGTGAGTCCAACTGGCGGACCTT 600
DB 5544 ACAGCCACCAAGCAAGCGAGCGCTGCTCCGTGTGAGTCCAACTGGCGGACCTT 5603
QY 601 GAGACTTCTGGGGGGAACAATGTGAACTTATCATGACGGGATACATTAAGTACAGGC 660
DB 5604 GAGACTTCTGGGGGGAACAATGTGAACTTATCATGACGGGATACATTAAGTACAGGC 5663
QY 661 TTGTGCACTGCGCTGGGGAATCCGCGATGATGATGATGATGATGATGATGATGATG 720
DB 5664 TTATCACTGCTGCTGGGGAATCCGCGATGATGATGATGATGATGATGATGATGATG 5723
QY 721 ACTAGCCGCTCACCAACCAATTAACCTCTGCTTAAATCTTGGGGGATGGGTAGCC 780
DB 5724 ACTAGCCGCTCACCAACCAATTAACCTCTGCTTAAATCTTGGGGGATGGGTAGCC 5783
QY 781 GCCCACTGCTCCCCAGAGTGTGCTTCACTTTGTAAGCGCGGATGCTGTGCTG 840
DB 5784 GCCCACTGCTCCCCAGAGTGTGCTTCACTTTGTAAGCGCGGATGCTGTGCTG 5843
QY 841 GCTGTGGCAGATTAAGGCTTGGGAGGTTGTTGATGATCTTGGCGGGATGAGACA 900
DB 5844 GCTGTGGCAGATTAAGGCTTGGGAGGTTGTTGATGATCTTGGCGGGATGAGACA 5903
QY 901 GAGAGTGGCAGCGCTGCTGCTTAAAGTCAATGAGCGGGAATGCTCCACGAG 960
DB 5904 GAGAGTGGCAGCGCTGCTGCTTAAAGTCAATGAGCGGGAATGCTCCACGAG 5963
QY 961 GACTGTTAACTTAATCTCTGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 5964 GACTGTTAACTTAATCTCTGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 6023
QY 1021 TGCGAGGATTAAGTGGTGGGATGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 6024 TGCGAGGATTAAGTGGTGGGATGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6083
QY 6084 CGGCTGATAGGCTTGGCTGCTGGGGGTAATATGTTTCCCAAGCACTATGCTGAG 6143
QY 1141 AGCGAGCGCGAGCGTGTCACTGATCTCTTCCGACTTATATCAACCACTGTTG 1200
DB 6144 AGCGAGCGCGAGCGTGTGTACTAGATCTCTTCCGACTTATATCAACCACTGCTG 6203
QY 1201 AAGAGGCTTCAACATGATTAACGAGGATGCTTCAAGCGCTGCTGGGCTGGGCTGA 1260
DB 6204 AAGAGGCTTCAACATGATTAACGAGGATGCTTCAAGCGCTGCTGGGCTGGGCTGA 6263
QY 1261 AGGAGTGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
DB 6264 AGGAGTGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6323

```


FT /product= "NS4a"
FT /note= "an isolated DNA comprising this sequence is
FT specifically claimed in Claim 1"
FT CDS 5544..5633
FT /tag= 8
FT /note= "an isolated DNA comprising this sequence is
FT specifically claimed in Claim 1"
FT CDS 5919..6371
FT /tag= 8
FT /product= "NS4b"
FT /note= "an isolated DNA comprising this sequence is
FT specifically claimed in Claim 1"
FT CDS 6372..9362
FT /tag= 8
FT /product= "NS5"
FT /note= "an isolated DNA comprising this sequence is
FT specifically claimed in Claim 1"
FT EP933426-A1.
PD 04-AUG-1999.
XX 28-DEC-1990; 99EP-00106005.
XX 25-JUN-1990; 90JP-00167466.
XX 31-AUG-1990; 90JP-00230921.
XX 09-NOV-1990; 90JP-00305605.
XX 28-DEC-1990; 90EP-00314371.
XX (OSAU) UNITV OSAKA.
PI Okayama H, Fuke I, Mori C, Takamizawa A, Yoshida I;
DR WPI; 1999-407152/35.
XX P-PSDB; AAY06423.
PT New hepatitis virus polypeptides, useful for diagnosing and treating
XX hepatitis infections.
PS Claim 1; Fig 2(1)-(16); 56pp; English.
XX This is the nucleotide sequence of non-A, non-B hepatitis virus (NANBV)
CC genomic cDNA containing the entire region of the open reading frame of
CC the NANBV genome. To obtain the genomic cDNA, NANBV RNAs were extracted
CC directly from NANBV particles contained in whole blood of a patient
CC having NANBV hepatitis, or from a resected liver of a patient having NANBV
CC hepatitis and liver cancer. The RNA was then converted to double-stranded
CC cDNA. A cDNA library was produced and screened using serum from a
CC convalescent patient having acute NANBV hepatitis and serum from a patient
CC having chronic NANBV hepatitis. The isolated cDNA allows recombinant
CC production of NANBV antigen polypeptides in microbial or eukaryotic cell
CC culture. The method provides the safe production of NANBV antigens with
CC high purity on a large scale at low cost without the biohazard associated
CC with multiplying virus in animals. Claimed NANBV nucleotide sequences are
CC useful for the diagnosis of NANBV hepatitis and for the recombinant
CC production of antigenic polypeptides. The polypeptides encoded by these
CC nucleotide sequences are useful as antigens for vaccines, and as
CC diagnostic reagents. (Updated on 20-MAR-2003 to correct PF field.)
CC (Updated on 20-MAR-2003 to correct PR field.)
XX
XX Sequence 9416 BP; 1905 A; 2826 C; 2678 G; 2007 T; 0 U; 0 Other;
SQ
Query Match 87.9%; Score 1796.6; DB 2; Length 9416;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY 1 TGGAGGCGCTTTCACAGGCTTCACCCAGTGATGCCACTTCTGTCCCAACAAG 60
DB TGGAGAGGTCTTCACAGGCTTCACCCATATAGATGACACTTCTGTCCCAACAAG 5063
QY 61 CAGGAGAGACAACTTCCCTTACCTGTGGGCTACAGGCTTACTGTGCGCTAGGGC 120
DB CAGGAGAGACAACTTCCCTTACCTGTGTAGATACAAAGCAAGGTGTGCGCGAGGCT 5123

QY 121 CAGGCCCACTTCATCATGGGATCAAAATGTGAAGTCTTCATACGGCTTAAAGCTACT 180
DB CAGGCCCACTTCATCATGGGATCAAAATGTGAAGTCTTCATACGGCTTAAAGCTACT 5183
QY 181 CTGCGCGGCGCAACACCTTCTGTATATAGCTGGGAGCGGCTCAAAACAGAGTACCCCTC 240
DB CTGCGCGGCGCAACACCTTCTGTATATAGCTGGGAGCGGCTCAAAACAGAGTACCCCTC 5243
QY 241 ACACACCCCATTAACAAATTCATCATGGCATGATGTCAGCGCATCTGGAGTGTCTACG 300
DB ACCACCCCATTAACAAATTCATCATGGCATGATGTCAGCGCATCTGGAGTGTCTACT 5303
QY 301 AGCACTGGGCTGTGTGGGCGGGGCTTTCAGCTCTGCTGCTATTTGCTTGACAACA 360
DB AGCACTGGGCTGTGTGGGCGGGGCTTTCAGCTCTGCTGCTGCTATTTGCTTGACAACA 5363
QY 361 GCGACGCTGTATTTGGGTTAGATCATCTTTGCCGGGCGGCGGCTATTTGCCGAGC 420
DB GCGAGTGTGTATTTGGGTTAGATCATCTTTGCCGGGCGGCGGCTATTTGCCGAGC 5423
QY 421 AGGGAAGTCTCTACAGAGATTTGATGATGAGATGAGATGCGGCTCGACCTCCCTTAC 480
DB AGGGAAGTCTCTCTACAGAGATTTGATGATGAGATGAGATGCGGCTCGACCTCCCTTAC 5483
QY 481 ATTCAGAGGGAATGACAGCTTCGCGAGCATTTCAAGCAAAAGCGCTTGGGTTGCTGAG 540
DB ATTCAGAGGGAATGACAGCTTCGCGAGCATTTCAAGCAAAAGCGCTTGGGTTGCTGCA 5543
QY 541 ACAAGCACAAGCAAGCGGAGCGGCTGCTCCGCTGGTGGAGTCCAAAGTGGCGAGCCCTT 600
DB ACAAGCACAAGCAAGCGGAGCGGCTGCTCCGCTGGTGGAGTCCAAAGTGGCGAGCCCTT 5544
QY 601 GAGACCTTCTGGGCGAAACACATGTGAACTTCATCAGCGGAGATACACTTACTTACAGGC 660
DB GAGACCTTCTGGGCGAAACACATGTGAACTTCATCAGCGGAGATACACTTACTTACAGGC 5663
QY 661 TTGTCACTTGTGCTGGGAAATCCGCGATTCGATCATGATGGCGTTTCAAGCTCTGTG 720
DB TTATCACTTGTGCTGGGAAATCCGCGATTCGATCATGATGGCGTTTCAAGCTCTTATC 5723
QY 721 ACTAGCCCGCTCACACCAATTAACCTCTGCTTACATCCGCGGAGTATGGGTACCC 780
DB ACTAGCCCGCTCACACCAATTAACCTCTGCTTACATCCGCGGAGTATGGGTACCC 5783
QY 781 GCCCAACTGCTCCCGCAGTGTGCTTCAAGCTTCTGTAGGCGCGCATTCCTGTGTCG 840
DB GCCCAACTGCTCCCGCAGTGTGCTTCAAGCTTCTGTAGGCGCGCATTCCTGTGTCG 5784
QY 841 GCTGTGGCAGCATAGGCTTGGGAAGTGTGTGACATCTTGGCGGCTTATGAGCA 900
DB GCTGTGGCAGCATAGGCTTGGGAAGTGTGTGACATCTTGGCGGCTTATGAGCA 5903
QY 901 GAGTGTGAGGCGGCTGTGCTTAAAGTCAAGACGGCGGAAATCCCTCCACCGAG 960
DB GAGTGTGAGGCGGCTGTGCTTAAAGTCAAGACGGCGGAAATCCCTCCACCGAG 5963
QY 961 GAGTGTGAGGCGGCTGTGCTTAAAGTCAAGACGGCGGAAATCCCTCCACCGAG 1020
DB GAGTGTGAGGCGGCTGTGCTTAAAGTCAAGACGGCGGAAATCCCTCCACCGAG 5964
QY 1021 TGGCAGGAGATATGCGTGGGACGTTGGGATGAGGAGGAGGAGGCTGTGAGTGAAC 1080
DB TGGCAGGAGATATGCGTGGGACGTTGGGATGAGGAGGAGGAGGCTGTGAGTGAAC 6083
QY 1081 CGGCTGATAGGCTTGTGCTGCGGGGTAAACATGTTTCCCAAGCAGCATATGTCAGAG 1140
DB CGGCTGATAGGCTTGTGCTGCGGGGTAAACATGTTTCCCAAGCAGCATATGTCAGAG 6143
QY 1141 AGCGAGCGCGAGCAGTGTCTAGATTCCTTCCGACCTTATCATCACTACAGTGTG 1200
DB AGCGAGCGCGAGCAGTGTCTAGATTCCTTCCGACCTTATCATCACTACAGTGTG 6144
QY 6144 AGCGAGCGCGAGCAGTGTCTAGATTCCTTCCGACCTTATCATCACTACAGTGTG 6203

```

QY 1201 AAGAGCTCCACAGATGATTAACGAGAGCTGCTCCAGCCTGCTCCGGCTGTTGAGCTA 1260
DB 6204 AAAAGGCTCCACAGATGATTAATGAAAGCTGCTCACAACCGGTTCCGGCTGTTGAGCTA 6263
QY 1261 AAGGATGTTTGGGATCTGGAATATGACAGTTTGGCTGACTTCAAGACTGGCTCCAGTCC 1320
DB 6264 AAGGATGTTTGGGATCTGGAATATGACAGTTTGGCTGACTTCAAGACTGGCTCCAGTCC 6323
QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTGATGCGCAAGTGGGTAAAGAGG 1380
DB 6324 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTGATGCGCAAGTGGGTAAAGAGG 6383
QY 1381 GTCTGCGGAGAGAGAGCATCATGACAGACCACTGCTCATGTGAGACAGATCAACGGA 1440
DB 6384 GTCTGCGGAGAGAGAGCATCATGACAGACCACTGCTCATGTGAGACAGATCAACGGA 6443
QY 1441 CATGTCAAAAAGGTTTCCATGAGATGTTGGGCTTAAGAAGCTGTATGATCATGTGGCAT 1500
DB 6444 CATGTCAAAAAGGTTTCCATGAGATGTTGGGCTTAAGAAGCTGTATGATCATGTGGCAT 6503
QY 1501 GGAACATTTCCCATCAAGCATACACAGAGGCGCTGACGCGCTCCAGGCGCAAC 1560
DB 6504 GGAACATTTCCCATCAAGCATACACAGAGGCGCTGACGCGCTCCAGGCGCAAC 6563
QY 1561 TATTCAGGCGCTGTGCGGCTGCTGCTGAGAGTACGTGAGGTTACGCGGCTGAGG 1620
DB 6564 TATTCAGGCGCTGTGCGGCTGCTGCTGAGAGTACGTGAGGTTACGCGGCTGAGG 6623
QY 1621 GATTTCCACTACGTGACAGCATGACCACTGACCAAGTAAATGCCCGTCCAGGTTCCA 1680
DB 6624 GATTTCCACTACGTGACAGCATGACCACTGACCAAGTAAATGCCCGTCCAGGTTCCG 6683
QY 1681 GCGCCGGAATTTCTCAAGAGTGAAGTGAGGCGGCTGACAGGATCGCTCCGCGTGC 1740
DB 6684 GCGCCGGAATTTCTCTGAGAGTGAAGTGAGGCGGCTGACAGGATCGCTCCGCGTGC 6743
QY 1741 AAACCTCTCTCAAGGAGAGGTCATCTTCAGGTGCGGCTCAACCAATACCTGTTGG 1800
DB 6744 AGGCTCTCTCAAGGAGAGGTCATCTTCAGGTGCGGCTCAACCAATACCTGTTGG 6803
QY 1801 TCGAGCTTCCATGAGAGCCCAACCGGATGATGACAGTGTCTCACTTCATGCTCACGAC 1860
DB 6804 TCGAGCTTCCATGAGAGCCCAACCGGATGATGACAGTGTCTCACTTCATGCTCACGAC 6863
QY 1861 CCCTCCCATCAACAGAGAGCGGTAAGCGGCGGCTGCGGAGGCGCTCCCGCTCC 1920
DB 6864 CCCTCCCATCAACAGAGAGCGGTAAGCGGCGGCTGCGGAGGCGCTCCCGCTCC 6923
QY 1921 TTGGCAGCTCTTCAAGTGAAGCTGTTGCGGCTTCTCGAAGGCGCATATTAAC 1980
DB 6924 TTGGCAGCTCTTCAAGTGAAGCTGTTGCGGCTTCTCGAAGGCGCATATTAAC 6983
QY 1981 CAAAATGACTTCCAGAGCGCTGACCTCATGAGGCGCACTTCCTGCGGCGATGAGATG 2040
DB 6984 CACCATGCTCTCCGAGCGCTGACCTCATGAGGCGCACTTCCTGCGGCGATGAGATG 7043
QY 2041 GGC 2043
DB 7044 GGC 7046

```

RESULT 11
AAD5321
AAD5321 standard: cDNA, 7987 BP.

XX AAD5321;
XX
XX 12-MAR-2002 (first entry)
XX
XX Hepatitis C virus (HCV) repliBartman/delta2U's cDNA.
XX
XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;

```

KM ss.
XX Hepatitis C virus.
OS
XX WO200189364-A2.
PN
XX 29-NOV-2001.
PD
XX 23-MAY-2001; 2001WO-US016822.
PF
XX 23-MAY-2000; 2000US-00576989.
PR
XX (UNITW ) UNITV WASHINGTON.
XX
XX Rice CM, Blight KJ;
PI
XX WPI; 2002-066755/09.
DR
XX Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
PS Claim 44; Page 66-69; 174pp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
XX polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX variants that have a transfection efficiency and ability to survive
XX subpassage greater than HCV that have wild-type polypeptide coding
XX regions. The polynucleotides of the invention are useful for identifying
XX a cell line that is permissive for infection with HCV and detecting
XX replication of HCV in cells of the cell line. They are also useful for
XX testing a compound for anti-viral properties and for inhibiting HCV
XX infection. They are also useful for the generation of defined HCV virus
XX stocks to develop in vitro and in vivo assays for virus neutralisation,
XX attachment, penetration and entry, structure/function studies on HCV
XX proteins and RNA elements and identification of new antiviral targets, a
XX systematic survey of cell culture systems and conditions to identify
XX those that support wild-type and variant HCV RNA replication and particle
XX release, production of adaptive HCV variants capable of more efficiency
XX replication in cell culture, production of HCV variants with altered
XX tissue or species tropism, establishment of alternative animal models for
XX inhibitor evaluation including those supporting HCV variant replication,
XX development of cell-free HCV replication assays, production of
XX immunogenic HCV particles for vaccination, engineering of attenuated HCV
XX derivatives as possible vaccine candidates, engineering of attenuated or
XX defective HCV derivatives for expression of heterologous gene products
XX for gene therapy and vaccine applications and for utilisation of the HCV
XX glycoproteins for targeted delivery of therapeutic agents to the liver
XX or other cell types with appropriate receptors. Vaccine comprising these
XX sequences is useful for inducing immunoprotection to HCV in a primate.
XX The present sequence is Hepatitis C virus (HCV) repliBartman/delta2U's
XX cDNA
SQ Sequence 7987 BP; 1647 A; 2368 C; 2243 G; 1729 T; 0 U; 0 Other;
XX
XX Query Match 87.5%; Score 1787; DB 6; Length 7987;
XX Best Local Similarity 92.2%; Pred. No. 0; Mismatches 0; Gaps 0;
XX Matches 1863; Conservative 0; Indels 0;
QY 1 TGGAGAGGCGTCTTCAAGAGGCTTCAAGGCTGATGAGGCTTCTGCTCCAAACAAAG 60
DB 3397 TGGAGAGGCGTCTTCAAGAGGCTTCAAGGCTGATGAGGCTTCTGCTCCAAACAAAG 3456
QY 61 CAGGCAAGAGACAACTTCCCTCACTGTTGGGCTGACAGGCTACTGTGTGGCTTGGGCT 120
DB 3457 CAGGCAAGAGACAACTTCCCTCACTGTTGGGCTGACAGGCTACTGTGTGGCTTGGGCT 3516
QY 121 CAGGCCCACTTCCATGAGGATCAAAATGAGAGTGTCTCATGAGGCTAAAGCTACT 180
DB 3517 CAGGCCCACTTCCATGAGGATCAAAATGAGAGTGTCTCATGAGGCTAAAGCTACTAG 3576
QY 181 CTGCGCGGCGCAACCTTGTGTATGAGTGTGGAGCGCTTCAAAACGAGGTCACTC 240

```

Db	357	TTGACGCGGGCAAGCCCTCTGTAATAGCTGGAGAGCCGTTCAAAACGAGGTTACTACC	363
Qy	241	ACAACCCCATTAACCAAAATTCATCATGGCATGCAATGTCAGCCACCTGAGAGTGTCAAG	300
Db	3637	ACACACCCCATTAACCAAAATCATCATGATGCAATGTCAGCTGACCTGAGAGTGTCAAG	3696
Qy	301	AGCACTTGGGAGCTGAGTGGGGGGGGTCCCTTGCAGCTCGCTGGGTATTTGTTACAACA	360
Db	3697	AGCACTTGGGAGCTGAGTGGGGGGGGTCCCTTGCAGCTCGCTGGGTATTTGTTACAACA	3756
Qy	361	GGCAGCGTGTCAATTGTGGGTAGGATCATCTTGTCCGGGCGGCGGGTATTTGCCGAC	420
Db	3757	GGCAGCGTGTCAATTGTGGGAGGATCATCTTGTCCGGGAAGCCGGCATATTTCCGAC	3816
Qy	421	AGGGAAGTCTCTACAGAGATTGGATGATGGAAGATGCGGTGCGACCTTCCCTTAC	480
Db	3817	AGGGAAGTCTCTTACCGGAGTTGGATGATGGAAGATGCGGTGCGCTCACCTCCCTTAC	3876
Qy	481	ATCCAGCAGGGAAATGACGTGCGCGAGAGTTCAAGCAAAAGGCTCGGGTTCCTCAG	540
Db	3877	ATCCAGCAGGGAAATGACGTGCGCGCAAAATTCAAACGAAGGCAATCGGGTTCCTCAA	3936
Qy	541	ACAGCCACCAAGCAAGCGAGGCGCTCTCCGCTGGAGAGTCAAGTGGCGAGCCCTT	600
Db	3937	ACAGCCACCAAGCAAGCGGAGGCTGCTCCCGTGGGAATCAAGTGGCGGACCTTC	3996
Qy	601	GAGACTTCTGCGCGGAAACACATGTGAACTTTCATCAGCGGGATACAGTACTTAGCAGCC	660
Db	3997	GAGCGCTTCTGGGCGAAGCATATGTGAAATTTTCATCAGCGGGATACAAATATTTAGCAGCC	4056
Qy	661	TTGTCCATCTGCGCTGGGAATCCGCGCATGCAATCAGATGAGGGCTCACAGCCTGTC	720
Db	4057	TTGTCCATCTGCGCTGGGAAACCCGCGATACATCAGATGAGGATTCACAGCCTTATC	4116
Qy	721	ACTAGCCCGCTCACCAACCAATCTACCTCTGCTTAACTCTGGGGGAGTGGGTAGCC	780
Db	4117	ACGACCCGCTCACCAACCAATACCTCTCTGTTTAACTCTGGGGGAGTGGGTAGCC	4176
Qy	781	GCCCAACTCGTCCCCCGCAGTGTGCTTCAGCTTTCGTAGGCGCGCGCATTTGCTGTGCG	840
Db	4177	GCCCAACTGTCCTCCCGCAGCGTGTTCGTGCTTTCGTAGGCGCGCGCATGCTGTGAGCG	4236
Qy	841	GCTGTGGCAGCATAGGCGCTTGGGAAGTGTCTTGTGACATCTTGGCGGGCTATGAGACA	900
Db	4237	GCTGTGGCAGCATAGGCGCTTGGGAAGTGTCTTGTGATATTTTGGCAGGTTATGAGACA	4296
Qy	901	GGAGTGGGAGGCGGCTGCTGGGCCCTTTAAGGCATGAGCGGCAAAATGCTCCACGAG	960
Db	4297	GGAGTGGGAGGCGGCTGCTGGGCCCTTTAAGGCATGAGCGGCGAATGCTCCACGAG	4356
Qy	961	GACTGTGTTAATTACTCCCTGCACTCTCTCTCTGTGTGCTGCTGTGTGCGGGATCGTG	1020
Db	4357	GACTGTGTTAATTACTCCCTGCTATCTCTCTCTCTCTGCGCGCTTATGTCGTGGGGTGTG	4416
Qy	1021	TGCGCAGCGATATCTGCGTGGGCACTGTGGTCCAGGGGAGGGGCGCTGTGCAATGATGAC	1080
Db	4417	TGCGCAGCGATATCTGCGTGGGCACTGTGGGCCAGGGGAGGGGCGCTGTGCAATGATGAC	4476
Qy	1081	CGGCTGATAGCGTTGCGCTCGCGGGGTAAACAATGTTCCCCACGCACTATGTGCGAG	1140
Db	4477	CGGCTGATAGCGTTGCTTGGGGGTAAACAAGTCTCCCGCAGCACTATGTGCTGAG	4536
Qy	1141	AGCGAGCGCGAGCAGTGTCACTAGATCCCTCGCATCTTATCATCAACCACTGTGTT	1200
Db	4537	AGCGAGCGCTGAGCACTGTGTCACTAGATCCCTCTTATCTTACATCACTCAAGTGTG	4596
Qy	1201	AAGAGGCTTCACAGTGAATTAAGAGGATGTCTCAACGCGCTGTCCGAGCTCGTGTGCTA	1260
Db	4597	AAGAGGCTTCACAGTGAATCAACAGAGACTGTCTCAAGCAGTGTCTCGAGCTCGTGTGCTA	4656
Qy	1261	AGGAGTGTTTGGGACTGTGATATGACAGTTTTGGCTGACTTCAAACTGTGCTCCAGTCC	1320
Db	4657	AGGAGTGTTTGGGATTTGATATGACAGGTTTGGACTGTATTCAAACCTGTGCTCCAGTCC	4716

OY	1321	AAAGTCTGCGCGGAGTTAACCGGAGAGTCCCTTTTTCATATCCAAAGTGGGTCAAGGGG	1380
Db	4717	AAAGTCTTGCGCGGAGTTACCGGAGAGTCCCTTTCTTCATATGTAACAAGTGGGTCAAGGGG	4776
OY	1381	GTCGCGGCGGAGACCGGATCATGACAGACCACTGTCATGTGGAGACAGATCACCGGA	1440
Db	4777	GTCGCGGCGGAGACCGGATCATGACAAACCACTGCCCATGTGGAGACAGATCACCGGA	4836
OY	1441	CATGTCAAAAAAGGTTCCATGAGGATCGTTGGGCTTAAGACCTGTAGTAACATGTGGCAT	1500
Db	4837	CATGTGAAAAAGGTTCCATGAGGATGTGGGGGCTGTAGACCTGTAGTAAACAAGTGGCAT	4896
OY	1501	GGAACTATCCCAATCAACGATTAACCAAGGGGCCCCGACAGGCCCCCAAGGCGCCAAAC	1560
Db	4897	GGAACTATCCCAATTAACGGGTAAACCAAGGGGCCCCCTGCAGGCCCTCCCGGCGCCAAAT	4956
OY	1561	TATTCACAGGCGCTGTGGCGGGGTGGCTGTAGAGAGTACGTGAGAGTTATACGCGGGTGGG	1620
Db	4957	TATTCAGAGGCGCTGTGGCGGGGTGGCTGTAGAGAGTACGTGAGAGTTATACGCGGGTGGG	5016
OY	1621	GATTTCCACTACGTGACGAGCATGACCACTGACACAAGTAAATATGCCCTGTCCAGGTTCCA	1680
Db	5017	GATTTCCACTACGTGACGAGCATGACCACTGACACAAGTAAAGTCCCGTCTCAGGTTCCG	5076
OY	1681	GGCCCCGAATTTCTTCAACAAGTGGATGGGGGTGGCTGACAGAGTACGCTCGGCGGTGC	1740
Db	5077	GGCCCCGAATTTCTTCAACAAGTGGATGGGGGTGGCTGACAGAGTACGCTCCAGCCTGC	5136
OY	1741	AAACCTCTCTTACCGGAGAGAGTCACTTCAGGTGGGGCTCAACAACATTAACCTGTGGTGGG	1800
Db	5137	AAACCTCTCTTACCGGAGAGAGTCACTTCCTGGTGGGGCTCAATCAATACCTGTGGTGGG	5196
OY	1801	TTCGACGCTCCCATGCGGAGCCCGGAACCGGAGTGAAGTGCATCTTCATGCTCACCGAC	1860
Db	5197	TTCACGCTCCCATGCGGAGCCCGGAACCGGAGTGAAGTGCATCTTCATGCTCACCGAC	5256
OY	1861	CCCTCCCAACATCAGACAGACAGACGCGTTAAACGCGACGAGCTGGCCACAGGGGAGTCTCCCTCC	1920
Db	5257	CCCTCCCAACATTAAGCGGAGAGACGCGTTAAAGCTGAAGCTGGCCACAGGGGATCTCCCTCC	5316
OY	1921	TTGGCCAGCTTCTTCACTGAGCTAAGCAAGTGTCTGGCGCTTCTCTGAAAGGCGACATACCTTAAC	1980
Db	5317	TTGGCCAGCTTCTTCACTGAGCTAAGCAAGTGTCTGGCGCTTCTTGAAGGCGACATGACCTAAC	5376
OY	1981	CAAAATACCTTCCCAAGCGCTGACCTGATGAGGCGCAACCTCGTGGCGGATGAGATG	2040
Db	5377	CGTCAATGCTTCCCGAGCGCTGACCTGATGAGGCGCAACCTCTGTGGCGGAGAGATG	5436
OY	2041	GGC 2043	
Db	5437	GGC 5439	
RESULT 12			
ID	AAA98968		
ID	AAA98968	standard; DNA; 7989 BP.	
XX	AAA98968;		
AC	08-FEB-2001	(first entry)	
XX	Hepatitis C virus DNA fragment SHQ ID NO: 4.		
DE	Cell culture; therapy; infection; vaccine; diagnosis; gene therapy; ds.		
XX	Hepatitis C virus.		
XX	DE19915178-A1.		
XX	05-OCT-2000.		
PD	03-APR-1999;	99DE-01015178.	
XX			

QY 1621 GATTTCACATAGTACGAGCATGACCACTGACCAAGTAAATGCCCGTCCAGGTTCCA 1680
DB 5017 GATTTCCACTAGTACGCGGCAATGACCACTGACCAAGTAAATGCCCGTCCAGGTTCCG 5076
QY 1681 GCGCCCGAATCTTTCACAGAAATGATGGGAGCGGCTGACAGAGTACGCTCCGGGTTCC 1740
DB 5077 GCGCCCGAATCTTTCACAGAAATGATGGGAGCGGCTGACAGAGTACGCTCCAGGTTCC 5136
QY 1741 AAACCTCTTCACGAGGAGAGTCACTTCAGTTCGGGCTGACCAATATCTGGTTCG 1800
DB 5137 AAACCTCTTCACGAGGAGAGTCACTTCAGTTCGGGCTGACCAATATCTGGTTCG 5196
QY 1801 TCGAGCTTCCTACGAGGAGAGTCACTTCAGTTCGGGCTGACCAATATCTGGTTCG 1860
DB 5197 TCGAGCTTCCTACGAGGAGAGTCACTTCAGTTCGGGCTGACCAATATCTGGTTCG 5256
QY 1861 CCGTCCCATGATGACGAGGAGAGTCACTTCAGTTCGGGCTGACCAATATCTGGTTCG 1920
DB 5257 CCGTCCCATGATGACGAGGAGAGTCACTTCAGTTCGGGCTGACCAATATCTGGTTCG 5316
QY 1921 TTGGCCAGCTCTTTCAGCTAGCCAGTTCCTGCGCTTCCTCGAAGGCGACATATACCTTACC 1980
DB 5317 TTGGCCAGCTCTTTCAGCTAGCCAGTTCCTGCGCTTCCTCGAAGGCGACATATACCTTACC 5376
QY 1981 CAAATGACTTCCGAGAGGCTGACCTCATTCGAGGCGCACTCTGTCGGGCGATGAGATG 2040
DB 5377 CGTCACTGACTCCCGAGAGGCTGACCTCATTCGAGGCGCACTCTGTCGGGCGAGAGATG 5436
QY 2041 GGC 2043
DB 5437 GGC 5439

RESULT 13

AAD25322
ID AAD25322 standard: cDNA, 7989 BP.

AC AAD25322;

DT 12-MAR-2002 (first entry)

DE Hepatitis C virus (HCV) repliBartman/Availi cDNA.

KM Hepatitis C virus; HCV; transfection; infection; virus neutralization;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KM ss.

OS Hepatitis C virus.

PH Key Location/Qualifiers

FT CDS 1801..7758

FT /tag= a

FT /product= "HCVrepliBartman polyprotein"

FT misc_feature 7766

FT /tag= b

FT /note= "Nucleotide creating Availi site"

PN WO200189364-A2.

PD 29-NOV-2001.

PE 23-MAY-2001; 2001WO-US016822.

PR 23-MAY-2000; 2000US-00576989.

RA (UNIM) UNIV WASHINGTON.

PI Rice CM, Blight KJ;

DR MPI; 2002-066755/09.

DR P-PSDB; AAE15717.

XX Hepatitis C virus variants having greater transfection efficiency and

PT ability to survive subpassage, useful as a vaccine for immunizing primate
FT to the virus, comprise non-naturally occurring viral sequences.

PS Claim 44; Page 69-71; 1749p; English.

CC The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) repliBartman/Availi cDNA

SQ Sequence 7989 BP; 1647 A; 2369 C; 2242 G; 1731 T; 0 U; 0 Other;

Query Match 87.5%; Score 1787; DB 6; Length 7989;

Best Local Similarity 92.2%; Pred. No. 0; Mismatches 160; Indels 0; Gaps 0;

Matches 1883; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 1 TGGGAGGCGTCTTTCACAGGCTGACCCAGTGCCTGACCTTCCTCCCAACAAG 60
DB 3397 TGGGAGGCGTCTTTCACAGGCTGACCCAGTGCCTGACCTTCCTCCCAACAAG 3456
QY 61 CAGGAGGAGCAACAATCTCCCTTACCTGTCGTCGTCACAGGCTGTCGCTAGGACC 120
DB 3457 CAGGAGGAGCAACAATCTCCCTTACCTGTCGTCGTCACAGGCTGTCGCTAGGACC 3516
QY 121 CAGGAGGAGCAACAATCTCCCTTACCTGTCGTCGTCACAGGCTGTCGCTAGGACC 180
DB 3517 CAGGAGGAGCAACAATCTCCCTTACCTGTCGTCGTCACAGGCTGTCGCTAGGACC 3576
QY 181 CTGCGCGGCGCAACACCTTCTCTGATAGGCTGGAGCCGTCACAAAGAGTCAACCTC 240
DB 3577 CTGCGCGGCGCAACACCTTCTCTGATAGGCTGGAGCCGTCACAAAGAGTCAACCTC 3636
QY 241 ACACACCCCAATACCAATTCATCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 300
DB 3637 ACACACCCCAATACCAATTCATCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 3696
QY 301 AGCACTTGGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 360
DB 3697 AGCACTTGGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3756
QY 361 GCGAGCGTGTGATTTGAGTATGATCTTGTCCGCGCGCGGCTATTTGCTCCGAC 420
DB 3757 GCGAGCGTGTGATTTGAGTATGATCTTGTCCGCGCGCGGCTATTTGCTCCGAC 3816
QY 421 AGGAAATCTCTTACAGAGTTCAGATGAGTGAAGAGTGCAGCTGACCTCCCTTAC 480
DB 3817 AGGAAATCTCTTACAGAGTTCAGATGAGTGAAGAGTGCAGCTGACCTCCCTTAC 3876
QY 481 ATCGAGCGGAGATTCAGCTGCGGAGAGTTCAGCAAAAAGCGCTGGGTTGCTGAG 540

```

Db 3877 ATCGAACAAGGATGAGCTCCGCAACATTCMAACAGAGGCAATCGGTTCTGCA 3936
Qy 541 ACAGCACCAGCAAGCGGAGCGCTGCTCCCGTGTGGAGTCCAAAGTGGAGCCCTT 600
Db 3937 ACAGCACCAGCAAGCGGAGCGCTGCTCCCGTGTGGAGTCCAAAGTGGAGCCCTT 3996
Qy 601 GAGACCTTCTGGGCGCAACACATGTGGAATTTCATCAGCGGATACATGACTTGAAGC 660
Db 3997 GAAGCTTCTGGGCGCAAGATGTGGAATTTTCATCAGCGGATACATGACTTGAAGC 4056
Qy 661 TTGTCACTCTGCTGGGAAATCCCGGATTCATCATGAGCGCTTCACAGCTCTGTG 720
Db 4057 TTGTCACTCTGCTGGGAAATCCCGGATTCATCATGAGCGCTTCACAGCTCTGTG 4116
Qy 721 ACTAGCCGCTCACCACCAATCTACCCCTGCTTAATCCTGCGGGGATGGGATG 780
Db 4117 ACCACCCGCTCACCACCAATCCTCTGCTTAATCCTGCGGGGATGGGATG 4176
Qy 781 GCCCACTGCTCCCCCAGTGTGCTGCTGAGCTTTGTAAGGCGCCGCAATGCTG 840
Db 4177 GCCCACTGCTCCCCCAGTGTGCTGCTGAGCTTTGTAAGGCGCCGCAATGCTG 4236
Qy 841 GCTGTGGCAGCATAGGCTTTGGAAGTGTGTTGACATCTTGCGGGCTATGAGCA 900
Db 4237 GCTGTGGCAGCATAGGCTTTGGAAGTGTGTTGATATTTTGGCAGTTATGAGCA 4296
Qy 901 GGAAGGCGAGGCGCGCTGCTGAGCTTTAGGTATGAGGCGGAAATGCCCTCCAC 960
Db 4297 GGAAGGCGAGGCGCGCTGCTGAGCTTTAGGTATGAGGCGGAAATGCCCTCCAC 4356
Qy 961 GACCTGTAACTTCTCCCTGCAATCTCTCTGCTGAGTCCCTGCTGCGGGTCTG 1020
Db 4357 GACCTGTAACTTCTCCCTGCAATCTCTCTGCTGAGTCCCTGCTGCGGGTCTG 4416
Qy 1021 TCGCAGCGCATACTGCTGCGCACTGTGGTCCAGGGAGGGGCTGTGCAATGAT 1080
Db 4417 TCGCAGCGCATACTGCTGCGCACTGTGGTCCAGGGAGGGGCTGTGCAATGAT 4476
Qy 1081 CGGCTGATAGCGCTTCCGCGGGGTAACATGTTTCCCGCAGCATATGCGCAG 1140
Db 4477 CGGCTGATAGCGCTTCCGCGGGGTAACATGTTTCCCGCAGCATATGCGCAG 4536
Qy 1141 AGCGAGCGCGAGCATGTGCTCATCATGATCTCTCCGACTTACTTACTTCACT 1200
Db 4537 AGCGAGCGCGAGCATGTGCTCATCATGATCTCTCCGACTTACTTACTTCACT 4596
Qy 1201 AAGAGGCTTCACAGTGAATTAACAGAGACTGTCTCCAGCTCGGCTGTGCT 1260
Db 4597 AAGAGGCTTCACAGTGAATTAACAGAGACTGTCTCCAGCTCGGCTGTGCT 4656
Qy 1261 AGGATGTTTGGGATCTGATATGCAAGTGTGGTGAATTCAGACCTGGCTCCAG 1320
Db 4657 AGGATGTTTGGGATCTGATATGCAAGTGTGGTGAATTCAGACCTGGCTCCAG 4716
Qy 1321 AAGCTCTGCGCGCATTAACCGGAGTCCCTTTTCTCATGCAAGTGGTATCAAG 1380
Db 4717 AAGCTCTGCGCGCATTAACCGGAGTCCCTTTTCTCATGCAAGTGGTATCAAG 4776
Qy 1381 GTCTGCGGCGAGCAGCATCATGAGACACCTGTGCTCATGTGAGCAGATCA 1440
Db 4777 GTCTGCGGCGAGCAGCATCATGAGACACCTGTGCTCATGTGAGCAGATCA 4836
Qy 1441 CATGTCAAAAACGGTTCCATGAGATCGTTGGGCTTAAGACTGTATGAATCAT 1500
Db 4837 CATGTCAAAAACGGTTCCATGAGATCGTTGGGCTTAAGACTGTATGAATCAT 4896
Qy 1501 GGAACATTTCCCATCAACGATACACAGGAGCCCTGACGCGCTCCCAAGCCAA 1560
Db 4897 GGAACATTTCCCATCAACGATACACAGGAGCCCTGACGCGCTCCCAAGCCAA 4956
Qy 1561 TATTCAGGCGCTGTGCGGGTGTGCTGTGAGAGATCATGTGAGGTTACGCGG 1620
Db 4957 TATTCAGGCGCTGTGCGGGTGTGCTGTGAGAGATCATGTGAGGTTACGCGG 5016

```

```

Qy 1621 GATTTCACTACGTAGCAGCATGACCACTGACAAAGTAAATGCCCGTGCAGTTCCA 1680
Db 5017 GATTTCACTACGTAGCAGCATGACCACTGACAAAGTAAATGCCCGTGCAGTTCCA 5076
Qy 1681 GCCCGCAATTTCTTCAACAGAGTGAATGGGCTGCGCTGCAAGTACGCTCCGCTGC 1740
Db 5077 GCCCGCAATTTCTTCAACAGAGTGAATGGGCTGCGCTGCAAGTACGCTCCGCTGC 5136
Qy 1741 AAACCTTCTTCAACAGAGTGAATGGGCTGCGCTGCAAGTACGCTCCGCTGC 1800
Db 5137 AAACCTTCTTCAACAGAGTGAATGGGCTGCGCTGCAAGTACGCTCCGCTGC 5196
Qy 1801 TCGAGCTCCATGAGCGAGCCGCAACCGAGTATGAGAGTGTCACTTCCATGCTCA 1860
Db 5197 TCGAGCTCCATGAGCGAGCCGCAACCGAGTATGAGAGTGTCACTTCCATGCTCA 5256
Qy 1861 CCTCCCATGAGCAGCAGACGAGCTTAAGCGAGGCTGCGCAGGGAGTCTCCCTCC 1920
Db 5257 CCTCCCATGAGCAGCAGACGAGCTTAAGCGAGGCTGCGCAGGGAGTCTCCCTCC 5316
Qy 1921 TTGGCCAGCTTTCAAGCTTACGATGTTGTGCGCTTCTCTGAAAGCCATATAC 1980
Db 5317 TTGGCCAGCTTTCAAGCTTACGATGTTGTGCGCTTCTCTGAAAGCCATATAC 5376
Qy 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCAACCTCTGTGCGGCGATGAG 2040
Db 5377 CCAATGACTTCCAGACGCTGACCTCATGAGGCAACCTCTGTGCGGCGATGAG 5436
Qy 2041 GGC 2043
Db 5437 GGC 5439

```

```

RESULT 14
ADJ57845
ID ADJ57845 strand; DNA; 7989 BP.
XX
AC ADJ57845;
XX
XX 06-MAY-2004 (first entry)
XX
DE HCV replicon encoding sequence.
XX
XX hepatitis C virus; HCV; Antinflammatory; Hepatotropic; Virucide; de;
XX HCV replicon.
XX
XX Unidentified.
XX
OS
XX
XX Key Location/Qualifiers
FH CDS 1801..7759
FT CDS /tag= a
FT CDS /product= "HCV replicon"
XX
PN MO2004015131-A2.
XX
PD 19-FEB-2004.
XX
XX 12-AUG-2003; 2003WO-US025260.
XX
XX PR 12-AUG-2002; 2002US-040261P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Gao M, Lemm JA, O'Boyle DR, Nower P, Rigat K, Sun J;
XX
XX MPI; 2004-180685/17.
XX
XX P-PSDB; ADJ57846.
XX
XX Use of hepatitis C virus assays or reporter assays, e.g. identifying a
XX compound that inhibits hepatitis C virus RNA replication or identifying a
XX compound that modulates the activity of a gene of interest.
XX

```


PS Claim 3, SEQ ID NO 1; 45bp; English.

XX The present invention relates to the use of hepatitis C virus (HCV)
CC assays for identifying a compound that inhibits HCV RNA replication and
CC reporter assays for identifying a compound that modulates the activity of
CC a gene of interest. The assays are useful for identifying a compound that
CC inhibits HCV RNA replication or for identifying a compound that modulates
CC the activity of a gene of interest. The HCV assay is useful for high
CC throughput screening that quantifies both the amount of HCV RNA
CC replication inhibitory activity associated with a test compound and the
CC amount of cytotoxicity associated with the test compound. The compound is
CC useful for treating hepatitis C infection. Assays of the invention have
CC distinct advantages when compared to qRT-PCR or other methods in that
CC assays of the invention may take place in situ in a detergent based crude
CC cell lysate, which requires no further preparation prior to performing
CC the assays. The assays do not also involve numerous manipulations to add
CC or subtract reagents after addition of test compounds and are desirably
CC based on a viral protein which is required by the HCV replicon for
CC replication. The present sequence represents a HCV replicon encoding
CC sequence used in the assay of the invention.
XX

Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other;

Query Match 87.5%; Score 1787; DB 12; Length 7989;

Best Local Similarity 92.2%; Pred.No. 0;

Matches 1883; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY	1	TGGGAGGGGCTTTCACAGGCTCAAGCCGATGAGTCCGCTGCTCCCAACAAG	60
DB	3397	TGGGAGGGGCTTTCACAGGCTCAAGCCGATGAGTCCGCTGCTCCCAACAAG	3456
QY	61	CAGGAGGAGACAACTTCCCTACCTGATGAGGCTACAGGCTACTGTGTGCGCTAAGGCC	120
DB	3457	CAGGAGGAGACAACTTCCCTACCTGATGAGGCTACAGGCTACTGTGTGCGCGCCAGGGCT	3516
QY	121	CAGGCGCCACCTTCATCATGAGTCAAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACT	180
DB	3517	CAGGCTCCACCTTCATCATGAGTCAAAATGTGGAAGTGTCTCATACGGCTAAAGCGTACG	3576
QY	181	CTGGCGGGGGCCAAACCCCTTGTGATAGAGCTGGGAGCGGCTCAAAACAGAGGTCAACCTC	240
DB	3577	CTGGCGGGGGCCAAACCCCTTGTGATAGAGCTGGGAGCGGCTCAAAACAGAGGTACTACCTC	3636
QY	241	ACACACCCCATTAACAAATTCATCATGAGTCAATGTGACGACCTGAGAGGTCTGTACG	300
DB	3637	ACACACCCCATTAACAAATTCATCATGAGTCAATGTGACGACCTGAGAGGTCTGTACG	3696
QY	301	AGCAGCTGGGTGCTGTGGGGGGGGCTTTCGACGCTTGTGCTGCTGATTTGCTTGACACA	360
DB	3697	AGCAGCTGGGTGCTGTGGGGGGGGCTTTCGACGCTTGTGCTGCTGATTTGCTTGACACA	3756
QY	361	GGCAGGCTGATTTGAGTGAATCATCTTGTCCGGGCGCGGCTATTGTTCCCGAC	420
DB	3757	GGCAGGCTGATTTGAGTGAATCATCTTGTCCGGGCGCGGCTATTGTTCCCGAC	3816
QY	421	AGGGAAGTCTCTACAGGAGTTCATGAGATGAGGAAGTGCAGCTGCGACCTCCCTTAC	480
DB	3817	AGGGAAGTCTCTTACCGGAGTTCATGAGATGAGGAAGTGCAGCTGCGACCTCCCTTAC	3876
QY	481	ATCGAGACGAGGAATGAGCTCCCGGACGCTTCAACGAAAAAGCGTGGGTTTCTGCGAG	540
DB	3877	ATCGAAGACGAGGAATGAGCTCCCGGACGCTTCAACGAAAAAGCGTGGGTTTCTGCGAA	3936
QY	541	ACAGCCACCAAGCAAGCGAGGCGCTGCTCCCGTGTGAGTCCAAAGTGGCGAGCGCTT	600
DB	3937	ACAGCCACCAAGCAAGCGAGGCGCTGCTCCCGTGTGAGTCCAAAGTGGCGAGCGCTT	3996
QY	601	GAGACCTTCTGGGCGGAACAATGTGGAATTCATCATGAGCGGAGTACATTAATTAAGCAGC	660
DB	3997	GAGACCTTCTGGGCGGAACAATGTGGAATTCATCATGAGCGGAGTACATTAATTAAGCAGC	4056
QY	661	TTGTCCACTCTGTGCTGGGAATCCCGGATTCATCATGAGTGGCGTTCAACGCTCTGTTC	720

DB	4057	TTGTCCACTCTGTGCTGGGAATCCCGGATTCATCATGAGTGGCGTTCAACGCTCTGTATC	4116
QY	721	ACTAGCCGCTACACCAACCAATTAACCTCTCTCTTAAATCTCTGGGGGATGGGTAGCC	780
DB	4117	ACTAGCCGCTACACCAACCAATTAACCTCTCTCTTAAATCTCTGGGGGATGGGTAGCC	4176
QY	781	GCCCACTCGTCCCCCAAGTGTGCTTCACTTTTGTGAAGCGCGGATGATGCTGTGCG	840
DB	4177	GCCCACTCGTCCCCCAAGTGTGCTTCACTTTTGTGAAGCGCGGATGATGCTGTGCG	4236
QY	841	GCTGTGGACACATAGGCTTGGGAAGTGTGATGACATCTTGGCGGGCTATGAGACA	900
DB	4237	GCTGTGGACACATAGGCTTGGGAAGTGTGATGACATCTTGGCGGGCTATGAGACA	4296
QY	901	GAGTGTGAGGCGCGCTTGTGCGCTTTAAGTCAATGAGCGCGAAATGCTCCACCGAG	960
DB	4297	GAGTGTGAGGCGCGCTTGTGCGCTTTAAGTCAATGAGCGCGAGATGCTCCACCGAG	4356
QY	961	GACCTGTTAATCTTAATCTCTGCAATCTCTCTGCTGCTGCTGTGCTGGGGTGTG	1020
DB	4357	GACCTGTTAATCTTAATCTCTGCAATCTCTCTGCTGCTGCTGTGCTGGGGTGTG	4416
QY	1021	TGGCAGGATTAAGCTGCGGACGATGAGTCCAGGGAGGGGCTGTGACATGATGAAC	1080
DB	4417	TGGCAGGATTAAGCTGCGGACGATGAGTCCAGGGAGGGGCTGTGACATGATGAAC	4476
QY	1081	CGGCTGATAGGCTTGTGCTGCGGGGTAAACATGTTTCCCAAGCACTATGTGCAAG	1140
DB	4477	CGGCTGATAGGCTTGTGCTGCGGGGTAAACATGTTTCCCAAGCACTATGTGCTAG	4536
QY	1141	AGCGACCGGACGACGATGATCAATCTCTTCCGACCTTAATCAATCAATCAATGTTG	1200
DB	4537	AGCGACCGGACGACGATGATCAATCTCTTCCGACCTTAATCAATCAATCAATGTTG	4596
QY	1201	AAAGAGCTTCAACGATGATTAACGAGAGTCTTCAAGCCCTGCTGCGGCTGAGCTA	1260
DB	4597	AAAGAGCTTCAACGATGATTAACGAGAGTCTTCAAGCCCTGCTGCGGCTGAGCTA	4656
QY	1261	AGGGAATGTTTGGGACTGATATGCAAGTGTGCTGCTTCAAGACCTGCTTCAAGTCC	1320
DB	4657	AGGGAATGTTTGGGATGATATGCAAGTGTGCTGCTTCAAGACCTGCTTCAAGTCC	4716
QY	1321	AAAGCTTCTGCGGCTTACCGGGAATCCCCCTTTTCTCATGCGCAAGGGGTGAGGAG	1380
DB	4717	AAAGCTTCTGCGGCTTACCGGGAATCCCCCTTTTCTCATGCGCAAGGGGTGAGGAG	4776
QY	1381	GTCTGGGGGAGAGGAGCATATGAGACCACTCTCATGTGAGGACAGATCAACCGGA	1440
DB	4777	GTCTGGGGGAGAGGAGCATATGAGACCACTCTCATGTGAGGAGAGATCAACCGGA	4836
QY	1441	CATGTCAAAAACGTTCTCATGAGATGTTGGGCTTAAAGCTGTGATGAATGTGGCAT	1500
DB	4837	CATGTCAAAAACGTTCTCATGAGATGTTGGGCTTAAAGCTGTGATGAATGTGGCAT	4896
QY	1501	GGAATATTCCTCATTAAGCATACACAGGGGCCCCCTGACGCGGCGCAAC	1560
DB	4897	GGAATATTCCTCATTAAGCATACACAGGGGCCCCCTGACGCGGCGCAAT	4956
QY	1561	TATTCAGGGGCTGTGGGCGGCTGTGAGAGTACGTGAGAGTTACGGGCTGTGGG	1620
DB	4957	TATTCAGGGGCTGTGGGCGGCTGTGAGAGTACGTGAGAGTTACGGGCTGTGGG	5016
QY	1621	GATTTCACTAGTACGAGCATGACATGACAACTGAAATGCGCTGCGAGGTTTCA	1680
DB	5017	GATTTCACTAGTACGCGGAGTACATGACAACTGAAATGCGCTGCTGAGGTTTCCG	5076
QY	1681	GCCCCGGAATTTTCAAGAGTGAATGGGCTGCGCTGCAAGGTAAGCTTCCGCGTGC	1740
DB	5077	GCCCCGGAATTTTCAAGAGTGAATGGGCTGCGCTGCAAGGTAAGCTTCCGCGTGC	5136
QY	1741	AAACCTTCTCAACGAGAGGATGACATTTCAAGTCTGAGGCTCAACAAATTAACCTGTGGG	1800
DB	5137	AAACCTTCTCAACGAGAGGATGACATTTCTGAGGCTCAACAAATTAACCTGTGGG	5196

QY 1801 TCGACAGCTCCATGCGAGCCGGAACCGAGTGTAGCAAGTGTCACTTCCATGTCCACCGAC 1860
 DB 5197 TCGAGCTCCATGCGAGCCGGAACCGAGTGTAGCAAGTGTCACTTCCATGTCCACCGAC 5256
 QY 1861 CCGTCCCATCATCAGCAGAGACGCGTAAAGCGAGCGTGGCCAGGGGGTCTCCCTCC 1920
 DB 5257 CCGTCCCATCATCAGCAGAGACGCGTAAAGCGAGCGTGGCCAGGGGGTCTCCCTCC 5316
 QY 1921 TTGGCAGCTTTCAGCTAGCCAGTGTCTGCGCCCTTCTCCAGAGGCGCATATAC 1980
 DB 5317 TTGGCAGCTTTCAGCTAGCCAGTGTCTGCGCCCTTCTCCAGAGGCGCATATAC 5376
 QY 1981 CAAATGACTTCCAGAGCGCTGCTCATGAGAGCCAACTCTGTGGCGCATGAGATG 2040
 DB 5377 CCGTCAATGCTCCCGGAGCGCTGCTCATGAGAGCCAACTCTGTGGCGCATGAGATG 5436
 QY 2041 GGC 2043
 DB 5437 GGC 5439
 RESULT 15
 AAL47276 standard; DNA; 7992 BP.
 ID AAL47276
 AC AAL47276;
 DT 30-AUG-2002 (first entry)
 DE Hepatitis C virus sub-genomic replicon clone 1377-NS3-3'UTR.
 XX Hepatitis; HCV; core-neo; NS3 proteinase/helicase; vaccine; diagnosis;
 KM virucide; hepatotropic; gene therapy; anti-viral; gene; ds.
 OS Hepatitis C virus.
 FT Key Location/Qualifiers
 FT CDS 342..1181
 FT /tag= a
 FT /product= "core-neo fusion protein"
 FT CDS 1801..7758
 FT /tag= b
 FT /product= "NS3 proteinase/helicase"
 PN WO200238793-A2.
 XX 16-MAY-2002.
 PD 02-NOV-2001; 2001WO-US046350.
 PF 07-NOV-2000; 2000US-0245866P.
 PR (ANAD-) ANADYS PHARM INC.
 PA Bichko V;
 PI WPI; 2002-490082/52.
 DR P-PSDB; AAO18000; AAO18001.
 XX Novel nucleic acid encoding replication competent recombinant hepatitis C
 PT virus genome useful for screening anti-hepatitis C virus therapeutics and
 PT for vaccine development.
 XX Claim 6; Page 43-47; 85pp; English.
 PS The present invention provides protein and coding sequences from
 CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and
 CC able to replicate efficiently when transfected into a susceptible cell
 CC line without reducing the growth rate of the cell line by more than 10
 CC fold. The sequences are useful for screening for anti-HCV therapeutics,
 CC for detecting antibodies to HCV in a biological sample such as blood,
 CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,

CC for deriving authentic HCV components such as replication-complement non-
 CC infectious, replication-defective infection-component, and replication-
 CC defective non-infectious HCV, in gene therapy or gene vaccination
 CC targeted to hepatic tissue for treating an animal infected or susceptible
 CC to HCV infection and for studying HCV infection and propagation. The
 CC present sequence is a clone of a fragment of the HCV genome which encodes
 CC the core-neo and NS3 proteinase/helicase proteins
 XX
 SQ Sequence 7992 BP; 1648 A; 2369 C; 2243 G; 1732 T; 0 U; 0 Other;
 Query Match 87.5%; Score 1787; DB 6; Length 7992;
 Best Local Similarity 92.2%; Pred. No. 0;
 Matches 1883; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
 QY 1 TGGGAGGGGCTTTCACAGGCGCTCACCGGAGATGCCACTTCCTGCCAAACAAG 60
 DB 3397 TGGGAGGGGCTTTCACAGGCGCTCACCGGAGATGCCACTTCCTGCCAAACAAG 3456
 QY 61 CAGCAGAGAGCAATCTCCCTTACCTGTGTGGGTACAGGCTTATGTGGCTAGGGCC 120
 DB 3457 CAGCAGAGAGCAATCTCCCTTACCTGTGTGGGTACAGGCTTATGTGGCTAGGGCC 3516
 QY 121 CAGGCCCCACTTCATCATGAGTAATGTGAAAGTGTCTCATACGCTTAAAGCTTACT 180
 DB 3517 CAGGCTCCACTTCATCATGAGTAATGTGAAAGTGTCTCATACGCTTAAAGCTTACG 3576
 QY 181 CTGGCGGGGCGCAACCCCTTGTGTATAGGCTGGGAGCGGCTCCAAAGAGGTCAACCC 240
 DB 3577 CTGGCGGGGCGCAACCCCTTGTGTATAGGCTGGGAGCGGCTCCAAAGAGGTCAACCC 3636
 QY 241 ACACACCCCATTAACCAATTCATCATGAGTATGACGAGCTGAGAGGTCTGTCAG 300
 DB 3637 ACACACCCCATTAACCAATTCATCATGAGTATGACGAGCTGAGAGGTCTGTCAG 3696
 QY 301 AGCACTGGGTCTGTGTGGCGGGGTCTTGCAGCTTGTGCTGCTTATGTTGACACA 360
 DB 3697 AGCACTGGGTCTGTGTGGCGGGGTCTTGCAGCTTGTGCTGCTTATGTTGACACA 3756
 QY 361 GGCAGCGTGTATGTGTGTAGATCATCTGTCCGGCGGCGCGCTTATTCCTCCGAC 420
 DB 3757 GGCAGCGTGTATGTGTGTAGATCATCTGTCCGGCGGCGCGCTTATTCCTCCGAC 3816
 QY 421 AGGGAAGTCTCTTACAGAGATTCATGATGATGAAAGTGTGCGGTCCCTCTTAC 480
 DB 3817 AGGGAAGTCTCTTACAGAGATTCATGATGATGAAAGTGTGCGGTCCCTCTTAC 3876
 QY 481 ATCGACAGGGAATGAGTGTGCGGACGATTCAGCAAAAGCGCTGGTGTGCGAG 540
 DB 3877 ATCGACAGGGAATGAGTGTGCGGACGATTCAGCAAAAGCGCTGGTGTGCGAG 5936
 QY 541 ACAGCCACCAACCAAGCGGAGCGCTGTCCCGTGTGTGGAATCCAGTGGGAGCGCTT 600
 DB 3937 ACAGCCACCAACCAAGCGGAGCGCTGTCCCGTGTGTGGAATCCAGTGGGAGCGCTT 3996
 QY 601 GAGACCTTGTGGCGGAAACACATGTGAACTTCATCAGCGGAGTACAGTCTTACAGAGC 660
 DB 3997 GAGACCTTGTGGCGGAAACACATGTGAACTTCATCAGCGGAGTACAGTCTTACAGAGC 4056
 QY 661 TTGTCCACTCTGCTGTGGATCCCGGATGTGATCATGATGAGCGGTTCACAGCTTGTCT 720
 DB 4057 TTGTCCACTCTGCTGTGGATCCCGGATGTGATCATGATGAGCGGTTCACAGCTTGTCT 4116
 QY 721 ACTAGCCGCTTACCAACCAATCTTACCTTCTGCTTAACTCTGTGGGGAATGGGTAGCC 780
 DB 4117 ACTAGCCGCTTACCAACCAATCTTACCTTCTGCTTAACTCTGTGGGGAATGGGTAGCC 4176
 QY 781 GCCCACTGCTCCCGGAGTGTGCTTGTGAGCGCGGCAATGCTGTGATG 840
 DB 4177 GCCCACTGCTCCCGGAGTGTGCTTGTGAGCGCGGCAATGCTGTGATG 4236
 QY 841 GCTGTGGAGCATAGGCTTGGGAAGTCTTGTGACATCTTGGCGGCTATGAGAGA 900
 DB 4237 GCTGTGGAGCATAGGCTTGGGAAGTCTTGTGACATCTTGTGAGATTTTGGCAGGTTATGAGAGA 4296

```
QY      1961  CAAATGACTTCCCAAGACGCTGACCTATCGAGGCCAACTCTGTGGCGCATGAGATG  2040
          |||||
Db      5377  CGTATCGCTCCCGAAGCTGACCTATCGAGGCCAACTCTGTGGCGCAGAGATG  5436
          |||||
QY      2041  GGC  2043
          ||||
Db      5437  GGC  5439

Search completed: February 25, 2005, 03:32:08
Job time : 964 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 03:10:04 ; Search time 366 Seconds
(without alignments)
9133.643 Million cell updates/sec

Title: US-09-664-363-20

Perfect score: 2043

Sequence: 1 TGGGAGGCGCTTCACAGC.....TGTGGCGCATGAGATGGCC 2043

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

Issued Patents, NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCITUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2043	100.0	2043	3	US-08-191-160-20 Sequence 20, Appl
2	2043	100.0	3750	1	US-08-191-160-22 Sequence 22, Appl
3	1796.6	87.9	7863	1	US-08-324-977-35 Sequence 35, Appl
4	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
5	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
6	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
7	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
8	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
9	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
10	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
11	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
12	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
13	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
14	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
15	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
16	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
17	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
18	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
19	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
20	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
21	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
22	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
23	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
24	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
25	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
26	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl
27	1796.6	87.9	7863	2	US-08-324-977-35 Sequence 35, Appl

28	1783.8	87.3	8638	4	US-10-029-907-6 Sequence 6, Appl
29	1783.8	87.3	8643	4	US-10-029-907-4 Sequence 4, Appl
30	1782.2	87.2	8638	4	US-10-029-907-7 Sequence 7, Appl
31	1782.2	87.2	8638	4	US-10-029-907-25 Sequence 25, Appl
32	1782.2	87.2	8648	4	US-10-029-907-5 Sequence 5, Appl
33	1780.6	87.2	8001	4	US-09-539-601-16 Sequence 16, Appl
34	1780.6	87.2	8001	4	US-09-539-601-22 Sequence 22, Appl
35	1780.6	87.2	11076	4	US-09-539-601-19 Sequence 19, Appl
36	1780.6	87.2	11076	4	US-09-539-601-25 Sequence 25, Appl
37	1777.4	87.0	8001	4	US-09-539-601-28 Sequence 28, Appl
38	1777.4	87.0	11076	4	US-09-539-601-31 Sequence 31, Appl
39	1756.6	86.0	9595	3	US-09-014-416-4 Sequence 4, Appl
40	1748.6	85.6	9413	4	US-09-827-688-6 Sequence 6, Appl
41	1719.8	84.2	9472	4	US-08-150-204E-96 Sequence 96, Appl
42	1345.4	65.9	9401	5	PCR-US91-02225-9 Sequence 9, Appl
43	1343.8	65.8	6785	3	US-08-444-818-65 Sequence 65, Appl
44	1343.8	65.8	8316	3	US-08-444-818-88 Sequence 88, Appl
45	1343.8	65.8	8987	3	US-08-444-818-137 Sequence 137, App

ALIGNMENTS

RESULT 1
US-08-191-160-20
; Sequence 20, Application US/08191160
; Patent No. 6210675
; GENERAL INFORMATION:
; APPLICANT: Highfield, Peter Edmund
; APPLICANT: Rodgers, Brian Collin
; APPLICANT: Tedder, Richard Seton
; APPLICANT: Barbara, John Anthony James
; TITLE OF INVENTION: Viral Agent
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Rothwell, Pigs, Ernst & Kurz
; STREET: 1700 K Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM AT compatible
; OPERATING SYSTEM: MS-DOS V3.2
; SOFTWARE: Wordperfect 5.0 (DOS text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/191.160
; FILING DATE:
; CLASSIFICATION:
; APPLICATION NUMBER: 07/628, 516
; FILING DATE: 17 DEC 1990
; APPLICATION NUMBER: UK 89 28 562.1
; FILING DATE: 18 DEC 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 414.0
; FILING DATE: 27 FEB 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 814.1
; FILING DATE: 03 MAR 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: E. Anthony Pigg
; REGISTRATION NUMBER: 27,195
; REFERENCE/DOCKET NUMBER: 1645-103A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 833-5740
; TELEFAX: (202) 833-5740
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2043 base pairs
; TYPE: nucleotide with corresponding protein
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
ORIGINAL SOURCE:
ORGANISM: human; serum infectious for PT-NANBH
IMMEDIATE SOURCE:
LIBRARY: clone 156/92
FEATURE:
LOCATION: from 1 to 2043 bp portion of the PT-NANBH
LOCATION: polyprotein
OTHER INFORMATION: probably encodes viral non-structural
OTHER INFORMATION: proteins
US-08-191-160-20

Query Match 100.0%; Score 2043; DB 3; Length 2043;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGAGGGGCTTTTCAAGGCTTCAAGCCGCTGAGTATGCTTCTGCTCCAAACAAAG 60
DB 1 TGGGAGGGGCTTTTCAAGGCTTCAAGCCGCTGAGTATGCTTCTGCTCCAAACAAAG 60
QY 61 CAGGAGAGAGCAACTTCCCTACTGAGGCTGACAGGCTACTGTGTGCGCTAGGGCC 120
DB 61 CAGGAGAGAGCAACTTCCCTACTGAGGCTGACAGGCTACTGTGTGCGCTAGGGCC 120
QY 121 CAGGCCCCCACTTCATCATGAGGATCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTTACT 180
DB 121 CAGGCCCCCACTTCATCATGAGGATCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTTACT 180
QY 181 CTGGCCGGGGCCCAACACCTTGTGTATAGAGCTGGAGCCGCTCCAAACAGAGTCAACCTTC 240
DB 181 CTGGCCGGGGCCCAACACCTTGTGTATAGAGCTGGAGCCGCTCCAAACAGAGTCAACCTTC 240
QY 241 ACACACCCCATTAACAAATTCATCATGAGGATGATGATGACAGGCTGAGAGGCTGTCACG 300
DB 241 ACACACCCCATTAACAAATTCATCATGAGGATGATGATGACAGGCTGAGAGGCTGTCACG 300
QY 301 AGCACCCTGGGTGCTGTGGGCGGGGCTCTTGACGCTTGGCTGTGATTTGTTGCAACA 360
DB 301 AGCACCCTGGGTGCTGTGGGCGGGGCTCTTGACGCTTGGCTGTGATTTGTTGCAACA 360
QY 361 GGCACCGTGTATGTGTGGTATGATCATCTTGTCCGGGCGGCGCTAATTTTCCGAC 420
DB 361 GGCACCGTGTATGTGTGGTATGATCATCTTGTCCGGGCGGCGCTAATTTTCCGAC 420
QY 421 AGGGAAGTCTCTACAGAGATTCAGATGATGAAAGATGAGCGCTGACCTCCCTTAC 480
DB 421 AGGGAAGTCTCTACAGAGATTCAGATGATGAAAGATGAGCGCTGACCTCCCTTAC 480
QY 481 ATCGACAGGAGATGCAAGTTCGCGAGCATTCAGCAAAAAGCGCTCGGTTGCTGCAG 540
DB 481 ATCGACAGGAGATGCAAGTTCGCGAGCATTCAGCAAAAAGCGCTCGGTTGCTGCAG 540
QY 541 AAGAGCCACCAACAGAGCGAGCGCTGTCTCCCGTGTGTGAGATCCAAAGTGGAGACCTT 600
DB 541 AAGAGCCACCAACAGAGCGAGCGCTGTCTCCCGTGTGTGAGATCCAAAGTGGAGACCTT 600
QY 601 GAGACCTTGTGGGCGAAACACATGTGGAATTCATCAGCGGGATACAGTACTTACAGAGC 660
DB 601 GAGACCTTGTGGGCGAAACACATGTGGAATTCATCAGCGGGATACAGTACTTACAGAGC 660
QY 661 TTGTCACTCTGCTGGGAATCCCGGATGTGATCATGTAGGCGTTCAAGCCTCTGTCT 720
DB 661 TTGTCACTCTGCTGGGAATCCCGGATGTGATCATGTAGGCGTTCAAGCCTCTGTCT 720
QY 721 ACTAGCCGCTCAACCAATTCATCTCTGCTTAAATCTCTGGGGGATGGGTAGGC 780
DB 721 ACTAGCCGCTCAACCAATTCATCTCTGCTTAAATCTCTGGGGGATGGGTAGGC 780
QY 781 GCCCAACTGCTCCCCCGAGTGTGCTTGTAGCTTTGTAGGCGCGGCAATTGCTGTGCG 840
DB 781 GCCCAACTGCTCCCCCGAGTGTGCTTGTAGCTTTGTAGGCGCGGCAATTGCTGTGCG 840

QY 841 GCTGTGGCAGCATAGGCTTGGGAAGGTGCTTGTGTGACATCTTGGCGGGCTATGAGACA 900
DB 841 GCTGTGTGACACATAGGCTTGGGAAGGTGCTTGTGTGACATCTTGGCGGGCTATGAGACA 900
QY 901 GAGATGAGAGCGGCTGTGGCTTTTAAAGTTCATGAGCGGAGAAATGCTCCACCGAG 960
DB 901 GAGATGAGAGCGGCTGTGGCTTTTAAAGTTCATGAGCGGAGAAATGCTCCACCGAG 960
QY 961 GACCTGTAACTTACTCCCTGCAATCTCTCTCTGTGTGCTGTGTGTGCTGTGCTGTG 1020
DB 961 GACCTGTAACTTACTCCCTGCAATCTCTCTCTGTGTGCTGTGTGTGCTGTGCTGTG 1020
QY 1021 TGGCAGAGATACAGTGTGCGGACAGTGTGTGCAAGGGAGGGGGCTGTGCAATGATGAC 1080
DB 1021 TGGCAGAGATACAGTGTGCGGACAGTGTGTGCAAGGGAGGGGGCTGTGCAATGATGAC 1080
QY 1081 CGGCTGATAGGCTTGCCTCGCGGGGTAAACCATGTTTCCCAAGCACTATGTGCGAGAG 1140
DB 1081 CGGCTGATAGGCTTGCCTCGCGGGGTAAACCATGTTTCCCAAGCACTATGTGCGAGAG 1140
QY 1141 AGCGACGGCGAGACAGTGTCACTCAGATCTCTCCGACTTATACCACTGATGCTGTG 1200
DB 1141 AGCGACGGCGAGACAGTGTCACTCAGATCTCTCCGACTTATACCACTGATGCTGTG 1200
QY 1201 AAGAGCTTCACCAATGATTTAAAGAGATGTCTCAAGCCCTGCTCGGGCTCGGGCTA 1260
DB 1201 AAGAGCTTCACCAATGATTTAAAGAGATGTCTCAAGCCCTGCTCGGGCTCGGGCTA 1260
QY 1261 AAGAGCTTCACCAATGATTTAAAGAGATGTCTCAAGCCCTGCTCGGGCTCGGGCTA 1260
DB 1261 AAGAGCTTCACCAATGATTTAAAGAGATGTCTCAAGCCCTGCTCGGGCTCGGGCTA 1260
QY 1281 AGGATGTTTGGAGCTGATATGCAAGTGTGCTGACTTCAAGACCTGTGCTCAGTCC 1320
DB 1281 AGGATGTTTGGAGCTGATATGCAAGTGTGCTGACTTCAAGACCTGTGCTCAGTCC 1320
QY 1321 AAGCTCTGCGCGATTAACGGGAGTCCCTTTTCTCATGCCAACTGTGGTAAAGGGG 1380
DB 1321 AAGCTCTGCGCGATTAACGGGAGTCCCTTTTCTCATGCCAACTGTGGTAAAGGGG 1380
QY 1381 GTCTGGCGGGAGACAGGATCATGACAGCACCTGTCTATGTGAGACACAGATCCGGA 1440
DB 1381 GTCTGGCGGGAGACAGGATCATGACAGCACCTGTCTATGTGAGACACAGATCCGGA 1440
QY 1441 CATGTCAAAAACGGTTCATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 CATGTCAAAAACGGTTCATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 GGAACATTCCTCCATCAACGATACACAGCGGCGCTGACAGCCTCCAGCGCAAAAC 1560
DB 1501 GGAACATTCCTCCATCAACGATACACAGCGGCGCTGACAGCCTCCAGCGCAAAAC 1560
QY 1561 TATTCAGAGGCGCTGTGGCGGGTGTGCTGAGAGATGCTGTGAGATTAACGGGGTGGGG 1620
DB 1561 TATTCAGAGGCGCTGTGGCGGGTGTGCTGAGAGATGCTGTGAGATTAACGGGGTGGGG 1620
QY 1621 GATTTCACTACGTGACAGATGACCACTGACCAAGTAAATGCGCTGACAGGTTCCA 1680
DB 1621 GATTTCACTACGTGACAGATGACCACTGACCAAGTAAATGCGCTGACAGGTTCCA 1680
QY 1681 GCCCCGGAATTCCTTCAAGAAAGTGTGAGGCTGCAAGGTGAGCTTCGCGGTGC 1740
DB 1681 GCCCCGGAATTCCTTCAAGAAAGTGTGAGGCTGCAAGGTGAGCTTCGCGGTGC 1740
QY 1741 AAACCTTCTCTACGGGAGAGATGATTCATTCAGAGTGTGGGCTTCAACATTAATCTGTTGGG 1800
DB 1741 AAACCTTCTCTACGGGAGAGATGATTCATTCAGAGTGTGGGCTTCAACATTAATCTGTTGGG 1800
QY 1801 TTGCAAGTCCCATGAGAGCGGGAACCGGATGAGAGAGTCACTTCCATGCTCACCGAC 1860
DB 1801 TTGCAAGTCCCATGAGAGCGGGAACCGGATGAGAGAGTCACTTCCATGCTCACCGAC 1860
QY 1861 CCCTCCACATCACAGAGAGAGCGCTAAGGCGAGGCTGCGCAAGGGGTCTCCCTCC 1920
DB 1861 CCCTCCACATCACAGAGAGAGCGCTAAGGCGAGGCTGCGCAAGGGGTCTCCCTCC 1920
QY 1921 TTGGCAGCTCTTCACTAGTACCAAGTGTGTGTGCGCTTCTCGAAGGCAATCAATTAC 1980

|||||
DB 1921 TTGGCCAGCTCTTCCAGTCCAGTGTCTGGCCCTTCCGAAAGCGCATATACCTACC 1980
QY 1981 CAAATGACTTCCCGAGCGCTGACCTCCATCGAGGCCAACCTCTGGGGGATGAGTAC 2040
DB 1981 CAAATGACTTCCCGAGCGCTGACCTCCATCGAGGCCAACCTCTGGGGGATGAGTAC 2040
QY 2041 GGC 2043
DB 2041 GGC 2043
RESULT 2
US-08-191-160-22
Sequence 22, Application US/08191160
Patent No. 6210675
GENERAL INFORMATION:
APPLICANT: Highfield, Peter Edmund
APPLICANT: Rodgers, Brian Colin
APPLICANT: Redder, Richard Seton
APPLICANT: Barbara, John Anthony James
TITLE OF INVENTION: Viral Agent
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSER: Rothwell, Figg, Ernst & Kurz
STREET: 1700 K Street
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
COMPUTER: IBM AT compatible
OPERATING SYSTEM: MS-DOS V3.2
SOFTWARE: Wordperfect 5.0 (DOS text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/191,160
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/628,516
FILING DATE: 17 DEC 1990
APPLICATION NUMBER: UK 89 28 562.1
FILING DATE: 18 DEC 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 414.0
FILING DATE: 27 FEB 1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 814.1
FILING DATE: 03 MAR 1990
ATTORNEY/AGENT INFORMATION:
NAME: E. Anthony Figg
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 1645-103A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 833-5740
TELEFAX: (202) 833-5744
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3750 base pairs
TYPE: nucleotide with corresponding protein
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
ORIGINAL SOURCE:
ORGANISM: human; serum infectious for PT-NANBH
IMMEDIATE SOURCE:
LIBRARY: cDNA clones from 3' end of the genome
FEATURE:
LOCATION: from 1 to 3750 bp portion of the PT-NANBH
OTHER INFORMATION: viral non-structural proteins
US-08-191-160-22

Query Match 100.0%; Score 2043; DB 3; Length 3750;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGGAGGGGCTTTTCAACAGGCTTCAACCAAGTGAATGCCACTTCTGTCACCAACAAAG 60
DB 1 TGGAGGGGCTTTTCAACAGGCTTCAACCAAGTGAATGCCACTTCTGTCACCAACAAAG 60
QY 61 CAGGAGAGAGAACTTCCCTACCTGATGGGCTACAGGCTGATGGGCTTGGGGC 120
DB 61 CAGGAGAGAGAACTTCCCTACCTGATGGGCTACAGGCTGATGGGCTTGGGGC 120
QY 121 CAGGCCCACTTCATCATGAGATCAATGAGTGTCTCATACGGCTAAAGCTTACT 180
DB 121 CAGGCCCACTTCATCATGAGATCAATGAGTGTCTCATACGGCTAAAGCTTACT 180
QY 181 CTGGCGGGGCCAAGACCTTGTGTATAGCTGGAGCCGTCCAAACGAGGTCACTC 240
DB 181 CTGGCGGGGCCAAGACCTTGTGTATAGCTGGAGCCGTCCAAACGAGGTCACTC 240
QY 241 ACACACCCCATPACCAATTCATCATGGCATGCAATGTCAGCCGACCTGAGGTCTACG 300
DB 241 ACACACCCCATPACCAATTCATCATGGCATGCAATGTCAGCCGACCTGAGGTCTACG 300
QY 301 AGCACCTGGGTGCTGTGGGGCGGGGCTCTTGACGCTCGGCTGATTTGTCACAA 360
DB 301 AGCACCTGGGTGCTGTGGGGCGGGGCTCTTGACGCTCGGCTGATTTGTCACAA 360
QY 361 GGCACCTGTCATTTGGGTAGATCATCTTGTCCGGCGCGGCTATTTGTTCCGAC 420
DB 361 GGCACCTGTCATTTGGGTAGATCATCTTGTCCGGCGCGGCTATTTGTTCCGAC 420
QY 421 AGGGAAGTCTTACCAAGAGTTCATGATGAGAAAGTGGCGCTCCCTTAC 480
DB 421 AGGGAAGTCTTACCAAGAGTTCATGATGAGAAAGTGGCGCTCCCTTAC 480
QY 481 ATCGAGAGGGGAATGACCTGCGGAGATTCAGCAAAAAGGCTGGGTGTCGAG 540
DB 481 ATCGAGAGGGGAATGACCTGCGGAGATTCAGCAAAAAGGCTGGGTGTCGAG 540
QY 541 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCGTGTGAGTCCAGTGGCGAGCCCTT 600
DB 541 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCGTGTGAGTCCAGTGGCGAGCCCTT 600
QY 601 GAGACCTTCTGGGGCAAAACATGTGAACTTCATCAGCGGATACATGACTTACAGGC 660
DB 601 GAGACCTTCTGGGGCAAAACATGTGAACTTCATCAGCGGATACATGACTTACAGGC 660
QY 661 TTGTCCACTGCGCTGGGAATCCCGGATGATGATGAGGCTTACAGGCTCTGTC 720
DB 661 TTGTCCACTGCGCTGGGAATCCCGGATGATGATGAGGCTTACAGGCTCTGTC 720
QY 721 ACTAGCCGCTCACCAACCAATCTACCTCTGCTTAAATCTTGGGGGATGAGTAC 780
DB 721 ACTAGCCGCTCACCAACCAATCTACCTCTGCTTAAATCTTGGGGGATGAGTAC 780
QY 781 GCCCAATCGCTCCCGGAGTGTCTTCAAGTTTGTAGGCGCGCATTTGCTGTG 840
DB 781 GCCCAATCGCTCCCGGAGTGTCTTCAAGTTTGTAGGCGCGCATTTGCTGTG 840
QY 841 GCTGTGGCAGATAGGCTTGGGAAGTGTGTGACATCTTGGCGGGCTATGAGACA 900
DB 841 GCTGTGGCAGATAGGCTTGGGAAGTGTGTGACATCTTGGCGGGCTATGAGACA 900
QY 901 GAGTGGCAGGCGCTGTGGCTTTAAGTCATGACGGGAAATGCCCTCCACGAG 960
DB 901 GAGTGGCAGGCGCTGTGGCTTTAAGTCATGACGGGAAATGCCCTCCACGAG 960
QY 961 GACCTGTAACTTACTCTTCACATCTCTCTCTGATGCGCTGATGCTGGGGTCTG 1020
DB 961 GACCTGTAACTTACTCTTCACATCTCTCTCTGATGCGCTGATGCTGGGGTCTG 1020

1021 TGGCGAGGATACGCTGCTGGGCACTGGGTCAGGGAGGGGGCTGTGCAATGATGAC 1080
1021 TGGCGAGGATACGCTGCTGGGCACTGGGTCAGGGAGGGGGCTGTGCAATGATGAC 1080
1081 CGGCTGATAGCGCTTGGCTCGGGGGGTAACATGTTTCCCGACCACTATGTGCAGAG 1140
1081 CGGCTGATAGCGCTTGGCTCGGGGGGTAACATGTTTCCCGACCACTATGTGCAGAG 1140
1141 AGCGAGCGCGAGCAGTGTCTCATCTGCTCTCCGACCTTACTATGACCCACCTGTG 1200
1141 AGCGAGCGCGAGCAGTGTCTCATCTGCTCTCCGACCTTACTATGACCCACCTGTG 1200
1201 AAGAGGCTTCACCACTGATTAACAGAGACTGCTCCAGCTGCTCGGCTGTGGCTA 1260
1201 AAGAGGCTTCACCACTGATTAACAGAGACTGCTCCAGCTGCTCGGCTGTGGCTA 1260
1261 AGGGATGTTGGGAGCTGATGATGCAAGTTTGGGCTGACTTCAAGCTGAGCTCCAGTCC 1320
1261 AGGGATGTTGGGAGCTGATGATGCAAGTTTGGGCTGACTTCAAGCTGAGCTCCAGTCC 1320
1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGTACAGGGG 1380
1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGTACAGGGG 1380
1381 GTCTGGCGGGAGACGGGATCATGCAAGACCACTGCTCATGTGGAGCAGATCACCGGA 1440
1381 GTCTGGCGGGAGACGGGATCATGCAAGACCACTGCTCATGTGGAGCAGATCACCGGA 1440
1441 CATGTCAAAAACGGTTCATGAGATCGTTGGGCTGAGACCTGATAGTAACTATGAGGAT 1500
1441 CATGTCAAAAACGGTTCATGAGATCGTTGGGCTGAGACCTGATAGTAACTATGAGGAT 1500
1501 GGAACATTTCCCATCAACGATACACAGGGGCGCTGACGCGCTCCCGACGCGCAAC 1560
1501 GGAACATTTCCCATCAACGATACACAGGGGCGCTGACGCGCTCCCGACGCGCAAC 1560
1561 TATTCAGGGGCGCTGTGGGCGGCTGCTGAGAGTATGAGAGGTTACGGGGTGGGG 1620
1561 TATTCAGGGGCGCTGTGGGCGGCTGCTGAGAGTATGAGAGGTTACGGGGTGGGG 1620
1621 GATTTCCACTAGTACGAGGATGACCACTGACCAAGTAAATGCCCGGCGAGGTTCCA 1680
1621 GATTTCCACTAGTACGAGGATGACCACTGACCAAGTAAATGCCCGGCGAGGTTCCA 1680
1681 GCGCCCGAATTTCTTCAAGAAATGGAGTGGGGTCCGCTGCAACAGTACCTCCGGGTGC 1740
1681 GCGCCCGAATTTCTTCAAGAAATGGAGTGGGGTCCGCTGCAACAGTACCTCCGGGTGC 1740
1741 AAACCTCTCTTACGAGAGAGGTCACTTCCAGTTCGGGCTCAACCAATACCTGTTGGG 1800
1741 AAACCTCTCTTACGAGAGAGGTCACTTCCAGTTCGGGCTCAACCAATACCTGTTGGG 1800
1801 TGGGAGCTTCCCATGAGAGCGGAGCCGAGTGTAGAGTCTCACTTCCATGCTCACCGAC 1860
1801 TGGGAGCTTCCCATGAGAGCGGAGCCGAGTGTAGAGTCTCACTTCCATGCTCACCGAC 1860
1861 CCTCCCATCATCAGAGCAGAGAGCGGTAAGCGGCTGAGGCGGAGCTTCCCGCTCC 1920
1861 CCTCCCATCATCAGAGCAGAGAGCGGTAAGCGGCTGAGGCGGAGCTTCCCGCTCC 1920
1921 TTGGCCAGCTTTCAGTACGACGATTTGTCTGCGCTTCTCGAAGGCGCATATACCT 1980
1921 TTGGCCAGCTTTCAGTACGACGATTTGTCTGCGCTTCTCGAAGGCGCATATACCT 1980
1981 CAAATGATTTCCCATGAGAGCGGTAAGCGGCTGAGGCGGAGCTTCTGAGGCGGATGAG 2040
1981 CAAATGATTTCCCATGAGAGCGGTAAGCGGCTGAGGCGGAGCTTCTGAGGCGGATGAG 2040
2041 GGC 2043
2041 GGC 2043

RESULT 3
US-08-324-977-35
Sequence 35, Application US/08324977
Patent No. 5747339
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Iba
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: Armstrong, Westerman, Hattori, McJeland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 7863 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7863
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..7863
OTHER INFORMATION: /note= "sequence = 1500 - 9362 of
OTHER INFORMATION: SEQ ID NO: 1"
US-08-324-977-35

Query Match 87.9%; Score 1796.6; DB 1; Length 7863;

Best Local Similarity 92.5%; Pred. No. 0; Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

1 TGGAGGGGGCTTCAACAGGCTCACCACGATGATGCCACTTCCGTCCCAACAAAG 60
3505 TGGAGAGTGTCTTACAGAGCTCACCACATATAGATGACACTTCTGTCCAGACAAAG 3564
61 CAGGAGAGAGCAACTTCCCTCACTGTGTGCGGTACAGGCTACTGTGTGCGGTAGAGGC 120
3565 CAGGAGAGAGCAACTTCCCTCACTGTGTGCGGTACAGGCTACTGTGTGCGGTAGAGGC 3624
121 CAGGCCCCACTTCATCATGTGGATCAAAATGTGAGTGTCTCATACGCTTAAAGCTTACT 180
3625 CAGGCCCCACTTCATCATGTGGATCAAAATGTGAGTGTCTCATACGCTTAAAGCTTACT 3684
181 CTGCGGGGGCCAAACCCCTTGTGTATAGCTGGGAGCCGTCCAAACAGAGTCAACCTTC 240
3685 CTGCAAGGGCCAAACCCCTTGTGTATAGCTGGGAGCCGTCCAAATGAGGTCAACCTTC 3744
241 ACACACCCCATTAACAAATTCATCATGTGATGATGATGATGATGATGATGATGATGATG 300
3745 ACCCACCCCATTAACAAATTCATCATGTGATGATGATGATGATGATGATGATGATGATG 3804
301 AGCACCCTGAGT 360
3805 ACACACTGT 3864
361 GGCAGAGT 420
3865 GGCAGAGT 3924
421 AGGGAAGTCTCTTACAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
3925 AGGGAAGTCTCTTACAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3984
481 ATGAGCAGAGAGATGTGAGT 540
3985 ATGAGCAGAGAGATGTGAGT 4044
541 ACAGCCCAAG 600
4045 ACAGCCCAAG 4104
601 GAGACCTTGT 660
4105 GAGACCTTGT 4164
661 TTGTCACTCTGT 720
4165 TTATCACTCTGT 4224
721 ACTAGCCGCTCAACCAATTCATCATGTGATGATGATGATGATGATGATGATGATGATG 780
4225 ACCAGCCGCTCAACCAATTCATCATGTGATGATGATGATGATGATGATGATGATGATG 4284
781 GCCCACTGT 840
4285 GCCCACTGT 4344
841 GCTGT 900
4345 GCTGT 4404
901 GAGAGT 960
4405 GAGAGT 4464
961 GACCTGT 1020
4465 GACCTGT 4524
1021 TGGCAGCGATTAATGT 1080

4525 TGTGACCAATTAATGT 4584
1081 CGAGT 1140
4585 CGAGT 4644
1141 AGCAGCGCGAG 1200
4645 AGCAGCGCGAG 4704
1201 AAGAGCTTCAACAG 1260
4705 AAGAGCTTCAACAG 4764
1261 AAGAGT 1320
4765 AAGAGT 4824
1321 AAGCTTGT 1380
4825 AAGCTTGT 4884
1381 GT 1440
4885 GT 4944
1441 CATGTCAAAAAGAGT 1500
4945 CATGTCAAAAAGAGT 5004
1501 GGAACATTTCCCATCAACCATCAACCATCAACCATCAACCATCAACCATCAACCATCAAC 1560
5005 GGAACATTTCCCATCAACCATCAACCATCAACCATCAACCATCAACCATCAACCATCAAC 5064
1561 TATTCAG 1620
5065 TATTCAG 5124
1621 GATTCACATTAATGT 1680
5125 GATTCACATTAATGT 5184
1681 GCCCCGGAATTTCTCAAGAGT 1740
5185 GCTCTGAATTTCTTCTGAGAGT 5244
1741 AAACCTTCTCTCAAGAGAGAGT 1800
5245 AGGCTTCTCTCAAGAGAGAGT 5304
1801 TGGCAGCTTCCATGAG 1860
5305 TCACAGCTTCAATGAG 5364
1861 CCTCCCATCAACAG 1920
5365 CCTCCCATCAACAG 5424
1921 TTGGCAGCTTCTCAAGT 1980
5425 TTGGCAGCTTCTCAAGT 5484
1981 CAATATGATTTCCAG 2040
5485 CACATATGATTTCCAG 5544
2041 GGC 2043
5545 GGC 5547

RESULT 4
US-08-384-616-35

/ Sequence 35, Application US/08384616
/ Patent No. 5847101
/ GENERAL INFORMATION:
/ APPLICANT: OKAYAMA, Hiroto
/ APPLICANT: FUKE, Isao
/ APPLICANT: MORI, Chisato
/ APPLICANT: TAKAMIZAWA, Akahisa
/ APPLICANT: YOSHIDA, Iwao
/ TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
/ TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
/ NUMBER OF SEQUENCES: 50
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Armstrong, Westerman, Hattori, McIeland &
/ ADDRESSEE: Naughton
/ STREET: 1725 K St. N.W. Suite 1000
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20006
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/384,616
/ FILING DATE:
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/769,996
/ FILING DATE: 02-OCT-1991
/ APPLICATION NUMBER: JP 2-167466
/ FILING DATE: 25-JUN-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-230921
/ FILING DATE: 31-AUG-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-305605
/ FILING DATE: 09-NOV-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/635,451
/ FILING DATE: 28-DEC-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stevens-Smith, Theresa M.
/ REGISTRATION NUMBER: 36,281
/ REFERENCE/DOCKET NUMBER: 900703B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 659-2930
/ TELEFAX: (202) 887-0357
/ TELEX: 440142
/ INFORMATION FOR SEQ ID NO: 35:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7863 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA from genomic RNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..7863
/ NAME/KEY: misc_feature
/ LOCATION: 1..7863
/ OTHER INFORMATION: /note= "sequence = 1500 - 9362 of
/ OTHER INFORMATION: SEQ ID NO: 1"
US-08-384-616-35
Query Match 87.9%; Score 1796.6; DB 2; Length 7863;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY 1 TGGAGAGGCTCTTACAGGCTTACCCACGATGAGTCCCACTTCTGTCCTCCAAACAAG 60
|||||

Db 3505 TGGAGAGTGTCTTACAGGCTTACCCATATAGATGACACTTCTGTCCAGACCAAG 3564
QY 61 CAGGACGAGAGCAACTTCCCTTACCTGTGGGATCCAGGCTACTGTGTGCTAGAGGCC 120
Db 3565 CAGGACGAGAGCAACTTCCCTTACCTGTAGATATCAAGCCACGGTGTGGCCAGGCT 3624
QY 121 CAGGCCCCACTTCCATCATGGGATTTAAATGTGGAAGTCTTCATACGGCTTAAAGCTCT 180
Db 3625 CAGGCCCCACTTCCATCATGGGATTTAAATGTGGAAGTCTTCATACGGCTTAAAGCTCT 3684
QY 181 CTGCGCGGACCAACACCCTTGCTGTATAGGCTGGAGCGGTCCAAACGAGGTACCCCTC 240
Db 3665 CTGACCGGACCAACACCCTTGCTGTATAGGCTGGAGCGGTCCAAACGAGGTACCCCTC 3744
QY 241 ACACACCCCATTAACCAATTCATCATGATGATCATGTACGCCGACTTGAAGTCTGTACG 300
Db 3745 ACCCACCCCATTAACCAATTCATCATGATGATCATGTACGCCGACTTGAAGTCTGTACT 3804
QY 301 AGCACCTGGTCTGTGTGGGCGGGTCTTTGACGCTGTGCTGTGCTTGTGTCACACA 360
Db 3805 AGCACCTGGTCTGTGTGGGCGGGTCTTTGACGCTGTGCTGTGCTTGTGTCACACA 3864
QY 361 GGCACGCTGTCTATGTGGGTAGATCATCTGTCCGGGCGGCGCTATTGTTCCGAC 420
Db 3865 GGCAGTGTGTCTATGTGGGTAGATCATCTGTCCGGGCGGCGCTATTGTTCCGAC 3924
QY 421 AGGGAAGTCTCTTACACAGAGTTCATGATGAAAGTGTGCGCTGCACCTCCCTTAC 480
Db 3925 AGGGAAGTCTCTTACACAGAGTTCATGATGAAAGTGTGCGCTGCACCTCCCTTAC 3984
QY 481 ATCGAGCAGGGAATTCAGTCTGCCGACGATTCAGCAAAAAGCGCTGGGTGTGAG 540
Db 3985 ATCGAGCAGGGAATTCAGTCTGCCGACGATTCAGCAAAAAGCGCTGGGTGTGAG 4044
QY 541 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCCGGTGTGAGTCCAAATGGCGAGCCCTT 600
Db 4045 ACAGCCACCAAGCAAGCGAGGCTGCTGCTCCCGGTGTGAGTCCAAATGGCGAGCCCTT 4104
QY 601 GAGACCTTCTGGGCGAACAATGTGAACTTCATCAGCGGATACAGTACTTACAGGC 660
Db 4105 GAGACCTTCTGGGCGAACAATGTGAACTTCATCAGCGGATACAGTACTTACAGGC 4164
QY 661 TTGTCACTCTGCTCTGGGAATCCCGCATTTGATACATGATGAGCGTTACAGCCCTGTTC 720
Db 4165 TTATCACTCTGCTCTGGGAATCCCGCATTTGATACATGATGAGCGTTACAGCCCTGTATC 4224
QY 721 ACTAGCCGCTACCAACCAATCTACCTCTGCTTAAATCATCTGGGGGATGAGTACGC 780
Db 4225 ACCAGCCGCTACCAACCAAGTACCTCTGCTTAAATCATCTGGGGGATGAGTACGC 4284
QY 781 GCCCAACTGCTCCCCAGATGCTGCTTCAAGCTTTCGTAGGCGCGCATTTGCTGTGCG 840
Db 4285 GCCCAACTGCTCCCCAGATGCTGCTTCAAGCTTTCGTAGGCGCGCATTTGCTGTGCG 4344
QY 841 GCTGTGGCAGATAGGCTTGGGAAGTGTCTTGTGACATTTGGCCGGGCTATAGACA 900
Db 4345 GCTGTGGCAGATAGGCTTGGGAAGTGTCTTGTGACATTTGGCCGGGCTATAGACA 4404
QY 901 GGAATGGCAGGCGGCTGCTGCTTAAAGTCATGACGCGGGAATGCGCTCCACGAG 960
Db 4405 GGAATGGCAGGCGGCTGCTGCTTAAAGTCATGACGCGGGAATGCGCTCCACGAG 4464
QY 961 GACCTGTAACTTACTCTGTCATCTCTCTCTGTGTCCTGTGCTGTGCGGTGCTG 1020
Db 4465 GACCTGTAACTTACTCTGTCATCTCTCTCTGTGTCCTGTGCTGTGCGGTGCTG 4524
QY 1021 TGGCAGGAGATCTGCTGTGCGACGTGGTCCAGGGGAGGGGCTGTGCAATGATGAAAC 1080
Db 4525 TGTGACAGAAATCTGCTGTGACACGTGGGTCCGGGAGAGGGGCTGTGAGATGAAAC 4584
QY 1081 CGGCTGATAGGTTTGGCGTGGGGGTAAACAATTTTCCCAAGCACTATGTGCAAG 1140
Db 4585 CGGCTGATAGGTTTGGCGTGGGGGTAAACAATTTTCCCAAGCACTATGTGCTGAG 4644


```

QY 1141 AGCGAGCGCGGAGAGAGTGTCTCAGATCTCTCCGACCTTACTATCAACCACTGTG 1200
    |||||
DB 4645 AGCGAGCGCGGAGAGAGTGTGTCTCAGATCTCTCCGACCTTACTATCAACCACTGTG 4704
QY 1201 AAGAGGCTCCACGAGTGTATTAACGAGAGCTGTCCAGCCCTGCTCCGAGTGTGCTTA 1260
    |||||
DB 4705 AAAAGGCTCCACGAGTGTATTAACGAGAGCTGTCCAGCCCTGCTCCGAGTGTGCTTA 4764
QY 1261 AGGGATGTTGGGAGCTGATATGACAGTGTGCTGACTTCAAGACCTGCTCCAGTCC 1320
    |||||
DB 4765 AGGGATGTTGGGAGCTGATATGACAGTGTGCTGACTTCAAGACCTGCTCCAGTCC 4824
QY 1321 AAGCTCTGCGCGGAGTTACCGGAGTCCCTTTTCTCATGCGCAAGTGGGTACAAGGGG 1380
    |||||
DB 4825 AAGCTCTGCGCGGAGTTACCTGAGTCCCTTTTCTCGTCCCAAGCGGGGTACAAGGGA 4884
QY 1381 GTCTGGCGGGGAGAGCGGATCATGACAGCACTGCTCATGTGAGCAGACAGATCACCGGA 1440
    |||||
DB 4885 GTCTGGCGGGGAGAGCGGATCATGACAGCACTGCTCATGTGAGCAGACAGATCACCGGA 4944
QY 1441 CATGTCAAAAAGGTTCCATGAGAGTCTGTGGGCTTAAGACCTGTATGAATCATGTGGCAT 1500
    |||||
DB 4945 CATGTCAAAAAGGTTCCATGAGAGTCTGTGGGCTTAAGACCTGTGAGCAACGTGGCAT 5004
QY 1501 GGAACATTCCTCCATCAAGCATACACAGAGGCGCCCTGACGCGCCCTCCAGGCGCAAC 1560
    |||||
DB 5005 GGAACATTCCTCCATCAAGCATACACAGAGGCGCCCTGACACCTCTTCAAGCGCCAAAC 5064
QY 1561 TATTCAGAGGCGCTGTGCGGGTGTGCTGTGAGAGTACGTGAGGTTACGCGGGTGGG 1620
    |||||
DB 5065 TATTCAGAGGCGCTGTGCGGGTGTGCTGTGAGAGTACGTGAGGTTACGCGGGTGGG 5124
QY 1621 GATTTCCACTAGTGAAGCATGACCACTGACACGTAATAATGCCCTGTCAGAGTTCCA 1680
    |||||
DB 5125 GATTTCCACTAGTGAAGCATGACCACTGACACGTAATAATGCCCTGTCAGAGTTCCG 5184
QY 1681 GCGCCGGAATTTCTCAAGAGTGAAGTGGGGTGGCGCTGACAGTACGCTCCGCGGTGC 1740
    |||||
DB 5185 GCGCCGGAATTTCTCTCGAGGTGAGCGGAGTGCAGTGTACAGGATGACCTCCGCGGTGC 5244
QY 1741 AAACCTCTCTCTAGGAGAGGTTCACATTCAGTGTGGGCTCAACCAATACCTGTTGGG 1800
    |||||
DB 5245 AGGCTCTCTCTAGGAGAGGTTCATTCAGTGTGGGCTCAACCAATACCTGTTGGG 5304
QY 1801 TCGCAGCTCCATGCGAGCCCGAACCAGATGTAAGCAGTGTCTCACTTCCATGCTACCGAC 1860
    |||||
DB 5305 TCACAGCTACATGCGAGCCCGAACCAGATGTAAGCAGTGTCTCACTTCCATGCTACCGAC 5364
QY 1861 CCTCCCATCATCAGCAGAGAGAGCGGTAGAGCGCAGGCTGCGAGGGGTCTCCCTCC 1920
    |||||
DB 5365 CCTCCCATCATCAGCAGAGAGAGCGGTAGAGCGCAGGCTGCGAGGGGTCTCCCTCC 5424
QY 1921 TTGGCAGGCTTTCAGTACGAGTGTGTGCGCTTCTCTGAGAGGCGACATATTAAC 1980
    |||||
DB 5425 TTGGCAGGCTTTCAGTACGAGTGTGTGCGCTTCTCTGAGAGGCGACATATTAAC 5484
QY 1981 CAAAATGACTTCCAGAGCGCTGACCTCATGAGGCGCAACCTCTGTGCGGCGATGAGATG 2040
    |||||
DB 5485 CACCATGTCTCTCGGAGCGCTGACCTCATGAGGCGCAACCTCTGTGCGGCGAGGATG 5544
QY 2041 GGC 2043
    |||||
DB 5545 GGC 5547
    |||||

```

```

RESULT 5
US-08-904-686A-35
; Sequence 35, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKE, Isao

```

```

; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McIeland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McIeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7863 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7863
; OTHER INFORMATION: /note="sequence = 1500 - 9362 of
; US-08-904-686A-35
; OTHER INFORMATION: SEQ ID NO: 1"
;
; Query Match 87.9%; Score 1796.6; DB 2; Length 7863;
; Best Local Similarity 92.5%; Pred. No. 0;
; Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```

```

QY 1 TGGAGGGCGGTCTTACAGAGGCTTACCGACGATGATGCCACTTCTGTCCAAACAAG 60
    |||||

```

Db 3505 TGGGAGGTGCTTTCACAGGCTTCAACCATATATAGACACTTCTTGTCACAGCAAG 3564
 QY 61 CAGGAGAGAGACAATTCCCTTACCTGTGGGTACAGAGCTACTGTGTGGCTAGGGCC 120
 Db 3565 CAGGAGAGAGACAATTCCCTTACCTGTGGGTACAGAGCTACTGTGTGGCTAGGGCC 3624
 QY 121 CAGGCCCCACCTTCATCATGGAGTCAAAATGTGAAAGTGTCTGATACGGCTAAAGCTTACT 180
 Db 3625 CAGGCCCCACCTTCATCATGGAGTCAAAATGTGAAAGTGTCTGATACGGCTAAAGCTTACT 3684
 QY 181 CTGGCGCGGCGCAACACCTTGTGTATAGCTGGGAGCGCTTCAAAACGAGTCAACCTTC 240
 Db 3685 CTGGACGGGCGCAACACCTTGTGTATAGCTGGGAGCGCTTCAAAATGAGTCAACCTTC 3744
 QY 241 ACACACCCCATATACAAATTCATCATGGAGTCAATGTGAGCTGTGAGCACTGTGAGCTGTCACG 300
 Db 3745 ACCACCCCATATACAAATTCATCATGGAGTCAATGTGAGCTGTGAGCACTGTGAGCTGTCACG 3804
 QY 301 AGCACCTGGGTGCTGTGGGCGGGGTCTTTCAGCTCTGGCTGTGCTGATTTGTTGACAAACA 360
 Db 3805 AGCACCTGGGTGCTGTGGGCGGGGTCTTTCAGCTCTGGCTGTGCTGATTTGTTGACAAACA 3864
 QY 361 GGCAGCGTGTATGTGTGGGTAGATCATCTTGTCCGGCGCGGCTTATTTGTTCCGAC 420
 Db 3865 GGCAGTGTGTATGTGTGGGTAGATCATCTTGTCCGGCGCGGCTTATTTGTTCCGAC 3924
 QY 421 AGGGAAGTCTCTTACAGAGATTCATGATGTAAGAGTGCAGCTGCACCTCTTAC 480
 Db 3925 AGGGAAGTCTCTTACAGAGATTCATGATGTAAGAGTGCAGCTGCACCTCTTAC 3984
 QY 481 ATCGACAGGAGATGACGCTGCGGAGAGCTTCAAGCAAAAACGCTGCGGTTCGTCAG 540
 Db 3985 ATCGACAGGAGATGACGCTGCGGAGAGCTTCAAGCAAAAACGCTGCGGTTCGTCAG 4044
 QY 541 AAGAGCCACCAACAGAGCGGAGCGCTGCTCCCGTGTGTGATGCCAAGTGGGAGCGCTT 600
 Db 4045 AAGAGCCACCAACAGAGCGGAGCGCTGCTCCCGTGTGTGATGCCAAGTGGGAGCGCTT 4104
 QY 601 GAGACCTTGTGGGCGAAACACATGTGGAATTCATCAGCGGAGTACAGACTTACAGAGC 660
 Db 4105 GAGACCTTGTGGGCGAAACACATGTGGAATTCATCAGCGGAGTACAGACTTACAGAGC 4164
 QY 661 TTGTCACTCTGCTGGGAATCCCGGATGTGATCATGATGAGCTTCAAGCTTCTGTC 720
 Db 4165 TTATCACTCTGCTGGGAATCCCGGATGTGATCATGATGAGCTTCAAGCTTCTATC 4224
 QY 721 ACTAGCCCCCTACCAACCAATTCATCTCTGCTTAAATCTCTGGGGGAGTGGGTAGCC 780
 Db 4225 ACCAGCCCCCTACCAACCAATTCATCTCTGCTTAAATCTCTGGGGGAGTGGGTAGCC 4284
 QY 781 GCCCAACTGCTCCCCAGTGTGCTGAGCTTTGTAGGCGCGGCAATGCTGTGTC 840
 Db 4285 GCCCAACTGCTCCCCAGTGTGCTGAGCTTTGTAGGCGCGGCAATGCTGTGTC 4344
 QY 841 GCTGTGGAGATAGAGCTTGGGAAGTGTGTGAGCATCTTGGCGGGCTATGAGAGA 900
 Db 4345 GCTGTGGAGATAGAGCTTGGGAAGTGTGTGAGCATCTTGGCGGGCTATGAGAGA 4404
 QY 901 GGAAGTGGAGGCGCGCTGCTGAGCTTTAAGTGTATAGAGGCGGAATGCTTCCACCGAG 960
 Db 4405 GGAAGTGGAGGCGCGCTGCTGAGCTTTAAGTGTATAGAGGCGGAATGCTTCCACCGAG 4464
 QY 961 GACCTGGTTAACTTACTCCCTGCACTCTGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTG 1020
 Db 4465 GACCTGGTTAACTTACTCCCTGCACTCTGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTG 4524
 QY 1021 TCGCAGCAGATATGCTGCTGCGCACGTGCTGCAAGGAGGAGGCTGTGAGTGGATGAAC 1080
 Db 4525 TGTGACGACATATGCTGCTGCGCACGTGCTGCAAGGAGGAGGCTGTGAGTGGATGAAC 4584
 QY 1081 CGGCTGATAGCTTGGCTGCTGCGGGGTAAACCATGTTTCCCGCAGCATATGTGCGAG 1140
 Db 4585 CGGCTGATAGCTTGGCTGCTGCGGGGTAAACCATGTTTCCCGCAGCATATGTGCGAG 4644

QY 1141 AGCGACGCGGAGACGCTGTCACTCAAGTCTTCCGACTTACTATCAACCACTGTG 1200
 Db 4645 AGCGACGCGGAGACGCTGTCACTCAAGTCTTCCGACTTACTATCAACCACTGTGCTG 4704
 QY 1201 AAGAGCTTCCACGATGATTTAAGAGATCTGTCCACGCGCTCTCCGGCTCGGGCTA 1260
 Db 4705 AAGAGCTTCCACGATGATTTAAGAGATCTGTCCACGCGCTCTCCGGCTCGGGCTA 4764
 QY 1261 AGGATGTTTGGAGCTGATATGACAGATTTTGGCTGACTTCAAGACTGTGCTCAGTTC 1320
 Db 4765 AGGATGTTTGGAGCTGATATGACAGATTTTGGCTGACTTCAAGACTGTGCTCAGTTC 4824
 QY 1321 AAGCTCTGCGCGATTAACGAGAGTCCCTTTTTCATGCCAACTGTGGTACAGGGG 1380
 Db 4825 AAGCTCTGCGCGAGTACCTGAGTCCCTTTTTCGTGTGCCAACTGTGGTACAGGGG 4884
 QY 1381 GTCTGGCGGGAGACGAGTATGACAGACCTGTCTCATGTGAGGACACATACCGGA 1440
 Db 4885 GTCTGGCGGGAGACGAGTATGACAGACCTGTCTCATGTGAGGACACATACCGGA 4944
 QY 1441 CATGTCAAAAACGCTTCATGAGATCGTTGGGCTTAAGACTGTATGATGTGAT 1500
 Db 4945 CATGTCAAAAACGCTTCATGAGATCGTTGGGCTTAAGACTGTATGATGTGAT 5004
 QY 1501 GGAACATTTCCCATCAACGATACACAGGAGCTTGCACGCTTCCCAAGGCTAAAC 1560
 Db 5005 GGAACATTTCCCATCAACGATACACAGGAGCTTGCACGCTTCCCAAGGCTAAAC 5064
 QY 1561 TATTTCAAGGCGCTGTGGCGGGGTGCTGTGAGATACGATGAGTAAACCGGGTGGG 1620
 Db 5065 TATTTCAAGGCGCTGTGGCGGGGTGCTGTGAGATACGATGAGTAAACCGGGTGGG 5124
 QY 1621 GATTTCCACTACGATGACGAGATGACACTGACCAAGTAATGCTGTCAGATTTCA 1680
 Db 5125 GATTTCCACTACGATGACGAGATGACACTGACCAAGTAATGCTGTCAGATTTCA 5184
 QY 1681 GCCCCGGAATTTCTTCAAGAAAGTGAATGGGTGCGGCTGCACAGTACGCTCCGCTGC 1740
 Db 5185 GCTCTGAATTTCTTCTCGAAGTGAACGAGTGTGCGGTTGCACAGGTAAGCTCCGCTGC 5244
 QY 1741 AAACCTCTCTTACGAGGAGAGATTCATTTCCAGTGTGGGCTTCAACCATATCTGTGG 1800
 Db 5245 AAGCTCTCTTACGAGGAGAGATTCATTTCCAGTGTGGGCTTCAACCATATCTGTGG 5304
 QY 1801 TCGAGCTTCCATGCGAGCGCGAACCAGATGAGAGTGTCACTTCCATGCTCACCGAC 1860
 Db 5305 TCGAGCTTCCATGCGAGCGCGAACCAGATGAGAGTGTCACTTCCATGCTCACCGAC 5364
 QY 1861 CCTTCCCATCAACAGAGAGCGGCTAAGCGAGCTGGCCAGGGGTCTTCCCTTCC 1920
 Db 5365 CCTTCCCATCAACAGAGAGCGGCTAAGCGAGCTGGCCAGGGGTCTTCCCTTCC 5424
 QY 1921 TTGGCCAGCTCTTCAAGCTTACGATGCTGTGTGCTGCTTCTGGAAGGCAATATAC 1980
 Db 5425 TTGGCCAGCTCTTCAAGCTTACGATGCTGTGTGCTGCTTCTTGAAGGCAATATAC 5484
 QY 1981 CAATATGACTTCCAGAGCGTGAACCTCATGAGGCAACTCTGTGCGGATGAGATG 2040
 Db 5485 CAATATGACTTCCAGAGCGTGAACCTCATGAGGCAACTCTGTGCGGATGAGATG 5544
 QY 2041 GGC 2043
 Db 5545 GGC 5547

RESULT 6
 US-09-315-850-35
 ; Sequence 35, Application US/09315850
 ; Patent No. 6217872
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKE, Isao

APPLICANT: MORI, Chiato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,850
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686
FILING DATE: 01-AUG-1997
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mcleland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 897-0357
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 7863 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7863
NAME/KEY: misc_feature
LOCATION: 1..7863
OTHER INFORMATION: /note= "sequence = 1500 - 9362 of
OTHER INFORMATION: SEQ ID NO: 1"
US-09-315-850-35

Query Match 87.9%; Score 1796.6; DB 3; Length 7863;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 1 TGGAGGGCGCTTTCACAGGCTCAACCACTGATGCCACTTCTCTGCCAACAAG 60
DB 3505 TGGAGAGGTCTTCAAGGCTCAACCAATAGATGACACTTCTGTGCCAGACAAG 3564
QY 61 CAGGACAGAGAACTTCCCCCTAAGTGGGTACAGGCTACTGTGGCCCTGGGGC 120
DB 3565 CAGGACAGAGAACTTCCCCCTAAGTGGGTACAGGCTACTGTGGCCCTGGGGC 3624
QY 121 CAGGCCCCCTCCATCATGGATCAAAATGTGAAGTCTCTATACGGCTAAAGCTACT 180
DB 3625 CAGGCCCCCTCCATCATGGATCAAAATGTGAAGTCTCTATACGGCTAAAGCTACT 3684
QY 181 CTGCGGGGCAACACCTTGTGTATAGCTGGAGCCGTCAAAACAGGTCAACCTTC 240
DB 3685 CTGCAAGGGCAACACCTTGTGTATAGCTGGAGCCGTCAAAACAGGTCAACCTTC 3744
QY 241 ACACACCCCTAACCAATTCTATGCGATGATGTACAGCCGACTGGAAGTCTGACG 300
DB 3745 ACCACCCCTAACCAATTCTATGCGATGATGTACAGCCGACTGGAAGTCTGACG 3804
QY 301 AGCACCCTGGGTCTGGTGGGCGGGGTCCTTGACGCTGGGCTGCGTATGTTGCAACA 360
DB 3805 AGCACCCTGGGTCTGGTGGGCGGGGTCCTTGACGCTGGGCTGCGTATGTTGCAACA 3864
QY 361 GGCAGCGTGTCTATGTGGGTAGATCATCTTGTCCGGCGGCGGCTATGTTCCGAC 420
DB 3865 GGCAGCGTGTCTATGTGGGTAGATCATCTTGTCCGGCGGCGGCTATGTTCCGAC 3924
QY 421 AGGGAATCTCTTACCAAGAGTTCATGATGATGAAAGTGGCGTCCGACTCTCCCTTAC 480
DB 3925 AGGGAATCTCTTACCAAGAGTTCATGATGATGAAAGTGGCGTCCGACTCTCCCTTAC 3984
QY 481 ATCGACAGGGAATGAGCTGCGGAGCAGTCAAGCAAAAAGCGCTGGGTTGTCAG 540
DB 3985 ATCGACAGGGAATGAGCTGCGGAGCAGTCAAGCAAAAAGCGCTGGGTTGTCAG 4044
QY 541 ACAGCCACCAACCAAGCGAGGCGCTGCTCCGTGTGGAGTCCAGTGGGAGCCCTT 600
DB 4045 ACAGCCACCAACCAAGCGAGGCGCTGCTCCGTGTGGAGTCCAGTGGGAGCCCTT 4104
QY 601 GAGACCTTGTGGGCGAACAATGTGAACTTCAATCAGCGGATACAGTACTTACGAGGC 660
DB 4105 GAGACCTTGTGGGCGAACAATGTGAACTTCAATCAGCGGATACAGTACTTACGAGGC 4164
QY 661 TTGTCCACTGCTCGGGAATCCCGGATGATGATGATGATGATGATGATGATGATGAT 720
DB 4165 TTATCCACTGCTCGGGAATCCCGGATGATGATGATGATGATGATGATGATGATGAT 4224
QY 721 ACTAGCCGCTCACACCAATCTACCTCTGCTTAAATCTCGGGGATGGGTAGCC 780
DB 4225 ACCAGCCGCTCACACCAATCTACCTCTGCTTAAATCTCGGGGATGGGTAGCC 4284
QY 781 GCCCACTGCTCCCCCAAGTCTGCTTCAAGTTTGTAGGCGCGGCACTTGTGTGCG 840
DB 4285 GCCCACTGCTCCCCCAAGTCTGCTTCAAGTTTGTAGGCGCGGCACTTGTGTGCG 4344
QY 841 GCTGTGGAGATAGGCTTGGGAAGTCTTGGAGATCTTGGCGGCTATGAGAGA 900
DB 4345 GCTGTGGAGATAGGCTTGGGAAGTCTTGGAGATCTTGGCGGCTATGAGAGA 4404
QY 901 GAGATGGCAGGCGCTCGTGGCTTAAAGTCAATGAGCGGGAATGCCCTCCACCGAG 960
DB 4405 GAGATGGCAGGCGCTCGTGGCTTAAAGTCAATGAGCGGGAATGCCCTCCACCGAG 4464
QY 961 GACTGTGTAATTAATCTCTCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 4465 GACTGTGTAATTAATCTCTCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 4524
QY 1021 TGGCAGAGATCTGCGTGGAGCGTGGGTTCCAGGGGAGGGGCTGTGAGTGAAG 1080
DB 4525 TGTGACGAAATCTGTGTGACACGTGGGTCCGGAGAGGGGCTGTGAGTGAAG 4584
QY 1081 CGGCTGATAGCGTGTGCGCTGCGGGGTAACATGTTTCCCGCACGACATATGTGCGAG 1140

```

Db      4585  CGGCTGATAGCGTTCCTCGGGGGATATGTTTCCCGACACATGTCCTAG 4644
Oy      1141  AGCGAGCGCGAGACGTCCTACAGATCTCTCCGACCTTACTATCACCAACTGTTG 1200
Db      4645  AGCGAGCGCGAGCGCGTTCCTACAGATCTCTCCGACCTTACTATCACCAACTGTTG 4704
Oy      1201  AAGAGGCTTCACCAATGATTAACAGAGACTCTCCACGCGCTGCTCGGCTGTTGCTA 1260
Db      4705  AAAAGGCTTCACCAATGATTAACAGAGACTCTCCACGCGCTGCTCGGCTGTTGCTA 4764
Oy      1261  AGGAGTGTGGGATGATGATGACAGTGTGGGCTGATTAAGACCTGGCTCCAGTCC 1320
Db      4765  AGGAGTGTGGGATGATGATGACAGTGTGGGCTGATTAAGACCTGGCTCCAGTCC 4824
Oy      1321  AAGCTCTGCGCGGATTAACGCGGAGTCCCTTTTCTCATGCGCAAGTGGTACAGGGG 1380
Db      4825  AAGCTCTGCGCGGAGTACCTGAGTCCCTTTTCTCGTGCACAGCGGGTACAGAGGA 4884
Oy      1381  GTCTGGCGGGAGACGGGATCATGACACCACTGCTCATGTGAGACACATCACCGGA 1440
Db      4885  GTCTGGCGGGAGACGGGATCATGACACCACTGCTCATGTGAGACACATCACCGGA 4944
Oy      1441  CATGTCAAAAAGGTTCCATGAGATCGTGGGCGCTTAAGACCTGTAAAGTAACTGGGAT 1500
Db      4945  CATGTCAAAAAGGTTCCATGAGATCGTGGGCGCTTAAGACCTGTAAAGTAACTGGGAT 5004
Oy      1501  GGAACATTTCCCATCAACGATACACCAAGGCGCGCTGACGCGCTCCCGACGCGCAAC 1560
Db      5005  GGAACATTTCCCATCAACGATACACCAAGGCGCGCTGACGCGCTCCCGACGCGCAAC 5064
Oy      1561  TATTCAGGCGGCTGTGGGGTGTGCTGAGAGATCGTGAAGTTACGCGGGTGGG 1620
Db      5065  TATTCAGGCGGCTGTGGGGTGTGCTGAGAGATCGTGAAGTTACGCGGGTGGG 5124
Oy      1621  GATTTCCACTAGTACGAGCATGACCACTGACCAAGTAAATGCCCGGACAGTTCCA 1680
Db      5125  GATTTCCACTAGTACGAGCATGACCACTGACCAAGTAAATGCCCGGACAGTTCCG 5184
Oy      1681  GCCCCGGAATTCCTACAGAAATGATGAGGGGTGCGGCTGACAGGTACGCTCCGGCTGC 1740
Db      5185  GCTCTGAATTTCTTCGAGAGGTGACGAGTGCCTGTCACAGTACGCTCCGGCTGC 5244
Oy      1741  AAACCTCTCTCAACGAGAGAGGTCAATTCAGGTGCGGCTCAACCAATTAAGTGTGGG 1800
Db      5245  AAGCCTCTCTCAACGAGAGAGGTCAATTCAGGTGCGGCTCAACCAATTAAGTGTGGG 5304
Oy      1801  TGGAGCTCCCATGCGAGCGCCGAAACCGGATGTAGAGTGTCTCACTGCTCAACCGAC 1860
Db      5305  TCACAGCTACATGCGAGCGCCGAAACCGGATGTAGAGTGTCTCACTGCTCAACCGAC 5364
Oy      1861  CCTCCCATCAACAGAGAGAGGTGAGCGAGGCTGAGCGAGGGGTCTCCCGCTGC 1920
Db      5365  CCTCCCATCAACAGAGAGAGGTGAGCGAGGCTGAGCGAGGGGTCTCCCGCTGC 5424
Oy      1921  TTGGCAGGCTTTCACTAGCAAGTGTCTGCGCTTCTCGAAGGCGCATATCAATTACC 1980
Db      5425  TTGGCAGGCTTTCACTAGCAAGTGTCTGCGCTTCTCGAAGGCGCATATCACTACC 5484
Oy      1981  CAAATGATTTCCCAAGCGTGAACCTCATGAGGCGCAACCTCTGTGGGGGATGAGAG 2040
Db      5485  CACCAATGTTCTCCGAGCGCTGACCTCATGAGGCGCAACCTCTGTGGGGGATGAGAG 5544
Oy      2041  GGC 2043
Db      5545  GGC 5547

```

```

; RESULT 7
; US-08-324-977-31
; Sequence 31, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:

```

```

; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKE, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESS: Armstrong, Westerman, Hatlori, McLealand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7917 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7862
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..7916
; OTHER INFORMATION: /note= "sequence = 1500 - 9416 of
; OTHER INFORMATION: SEQ ID NO: 1"
; US-08-324-977-31

```

```

Query Match      87.9%; Score 1796.6; DB 1; Length 7917;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
Oy      1  TTGGAGGCGCTTTACAGCGCTTACCAAGTGAATGCGCACTTCTGTCCTGCAACCAAG 60

```

Db 3505 TGGGAGAGTGTCTTCAAGGCTTCAACCATATAGATGACAATTCTTGTCTCCAGACCAAG 3564
 QY 61 CAGGAGAGAGCAACTTCCCTTACCTGAGGCTATCAGGCTATCTGTGCGCTAGAGGC 120
 Db 3565 CAGGAGAGAGCAACTTCCCTTACCTGAGGCTATCAGGCTATCTGTGCGCTAGAGGC 3624
 QY 121 CAGGCCCCACCTCATCATGAGATCAAAATGTGGAAGTGTCTCATACAGGCTAAAGCTTACT 180
 Db 3625 CAGGCCCCACCTCATCATGAGATCAAAATGTGGAAGTGTCTCATACAGGCTAAAGCTTACT 3684
 QY 181 CTGCGCGGCGCAACACCTTCTGTATAGGCTGGAGCGCTCCAAAAGAGGTCAACCTTC 240
 Db 3695 CTGCAACGGGCAACACCTTCTGTATAGGCTGGAGCGCTCCAGATATAGGTCAACCTTC 3744
 QY 241 ACACACCCCATATACAAATTCATCATGAGATGATGATGACGCGACCTGGAGGTCGTCAG 300
 Db 3745 ACCACCCCATATACAAATTCATCATGAGATGATGATGACGCGACCTGGAGGTCGTCAG 3804
 QY 301 AGCAGCTGGGTCTGGTGGGCGGGGTCTTTCAGAGCTCTGGCTGTATTGTTCACACA 360
 Db 3805 AGCAGCTGGGTCTGGTGGGCGGGGTCTTTCAGAGCTCTGGCTGTATTGTTCACACA 3864
 QY 361 GCGAGCGTGTATGTGTGGTATGATCATCTTGTCCGGCGGCGGCTATTTGTTCCGAC 420
 Db 3865 GCGAGTGTGTATGTGTGGTATGATCATCTTGTCCGGCGGCGGCTATTTGTTCCGAC 3924
 QY 421 AGGGAAGTCTCTACAGAGATTCAGATGAGATGGAAGTGGCGGTGAGACCTCCCTTAC 480
 Db 3925 AGGGAAGTCTCTACAGAGATTCAGATGAGATGGAAGTGGCGGTGAGACCTCCCTTAC 3984
 QY 481 ATCGACAGAGGATGAGCTCGCCGAGCAGTTCAAGCAAAAAGCGCTCGGTTGCTGAC 540
 Db 3985 ATCGACAGAGGATGAGCTCGCCGAGCAGTTCAAGCAAAAAGCGCTCGGTTGCTGAC 4044
 QY 541 ACAGCCCAAGCAAGCGAGCGCTGTCTCCGTGTGTGAGTCCAGTGGCGAGCCTT 600
 Db 4045 ACAGCCCAAGCAAGCGAGCGCTGTCTCCGTGTGTGAGTCCAGTGGCGAGCCTT 4104
 QY 601 GAGACCTTCTGGGCGAAACATGTGGAATTCATCAGGCGGATACAGTACTTACAGAGC 660
 Db 4105 GAGACCTTCTGGGCGAAACATGTGGAATTCATCAGGCGGATACAGTACTTACAGAGC 4164
 QY 661 TTGTCACTCTGCTGGGAATCCCGGATGTGATCATCTGATGGCGTTCAAGCCTTGTCT 720
 Db 4165 TTGTCACTCTGCTGGGAATCCCGGATGTGATCATCTGATGGCGTTCAAGCCTTGTCT 4224
 QY 721 ACTAGCCGCTCACCAACCAATCTACCTCTCTGTTTAAATCTCTGGGGGATAGTGGC 780
 Db 4225 ACTAGCCGCTCACCAACCAATCTACCTCTCTGTTTAAATCTCTGGGGGATAGTGGC 4284
 QY 781 GCCCAATCTGCTCCCGCAAGTGTGTTAGCTTTCGTAGGCGCGGCAATTGCTGTGG 840
 Db 4285 GCCCAATCTGCTCCCGCAAGTGTGTTAGCTTTCGTAGGCGCGGCAATTGCTGTGG 4344
 QY 841 GCTGTGGAGCATAGGCTTGGGAAAGTGTGTGACATCTTGGCGGGCTATAGAGA 900
 Db 4345 GCTGTGGAGCATAGGCTTGGGAAAGTGTGTGACATCTTGGCGGGCTATAGAGA 4404
 QY 901 GGAAGTGGAGGCGCGCTGTGGCTTTTAAAGTATAGAGCGGCGAAATGCTTCCACCGAG 960
 Db 4405 GGAAGTGGAGGCGCGCTGTGGCTTTTAAAGTATAGAGCGGCGAAATGCTTCCACCGAG 4464
 QY 961 GACCTGTGTTAACTTACTCCCTGCAATCTCTCTCTGTGGTCCCTGATCGGGGTCGTG 1020
 Db 4465 GACCTGTGTTAACTTACTCCCTGCAATCTCTCTCTGTGGTCCCTGATCGGGGTCGTG 4524
 QY 1021 TCGCAGCGATATCTAGCTGAGACGAGGTTCAGAGGAGGAGGAGCTGTGAGTATGAAAC 1080
 Db 4525 TCGCAGCGATATCTAGCTGAGACGAGGTTCAGAGGAGGAGGAGCTGTGAGTATGAAAC 4584
 QY 1081 CGGCTGATAGCTTGGCTCTCGCGGGGTAAACATGTTTTCCCAACGCACTATGTGCCAG 1140
 Db 4585 CGGCTGATAGCTTGGCTCTCGCGGGGTAAACATGTTTTCCCAACGCACTATGTGCCAG 4644

QY 1141 AGCGAGCGCGAGCAGTGTCACTAGATCTCTCCGACTTACTATACCAACTGTG 1200
 Db 4645 AGCGAGCGCGAGCAGCGGTCTTACTAGATCTCTCCAGCTTACATGACTGAGCTG 4704
 QY 1201 AAGAGCTTCAACAGTGTATTAAGAGAGCTGCTCAAGCCTGTCTCGGCTGTGGCTA 1260
 Db 4705 AAGAGCTTCAACAGTGTATTAAGAGAGCTGCTCAACCGTGTCTCGGCTGTGGCTA 4764
 QY 1261 AAGAGTGTGAGACTGTATAGCAAGTGTGAGCTTGAACCTGAGCTGAGTCC 1320
 Db 4765 AAGAGTGTGAGACTGTATAGCAAGTGTGAGCTTGAACCTGAGCTGAGTCC 4824
 QY 1321 AAGCTCTGCGCGCAGTACCGGAGTCCCTTTTCTATGCAAGCTGTGAGTAAAGGG 1380
 Db 4825 AAGCTCTGCGCGCAGTACCGGAGTCCCTTTTCTATGCAAGCTGTGAGTAAAGGG 4884
 QY 1381 GTCTGGCGGGAGAGCGGATCATGAGACCACTGCTCATGTGAGCAGATCAACCGGA 1440
 Db 4885 GTCTGGCGGGAGAGCGGATCATGAGACCACTGCTCATGTGAGCAGATCAACCGGA 4944
 QY 1441 CATGTCAAAAAGGTTCCATGAGATGTGTGGGCTTAAGCCTGTATGATATGTGGAT 1500
 Db 4945 CATGTCAAAAAGGTTCCATGAGATGTGTGGGCTTAAGCCTGTATGATATGTGGAT 5004
 QY 1501 GGAACATTCGCCATCAACGATACACAGAGGCGCCCTGACGCGCTCCAGCGCAAAAC 1560
 Db 5005 GGAACATTCGCCATCAACGATACACAGAGGCGCCCTGACGCGCTCCAGCGCAAAAC 5064
 QY 1561 TATTCAGAGCGCTGTGGCGGGTGTCTGTAGAGATGATGAGAGTTACCGGGGTGGG 1620
 Db 5065 TATTCAGAGCGCTGTGGCGGGTGTCTGTAGAGATGATGAGAGTTACCGGGGTGGG 5124
 QY 1621 GATTTTCACTAGTACAGCATGACCACTGACCAAGTAAATGCCCCGCTGAGGTTCCA 1680
 Db 5125 GATTTTCACTAGTACAGCATGACCACTGACCAAGTAAATGCCCCGCTGAGGTTCCG 5184
 QY 1681 GCCCGCAATTCCTACAGAAAGTGTAGTGGGTTGGCGCTGACAGGTAAGCTCCGCGTGC 1740
 Db 5185 GCCCGCAATTCCTACAGAAAGTGTAGTGGGTTGGCGCTGACAGGTAAGCTCCGCGTGC 5244
 QY 1741 AAACCTCTCTTACGAGGAGAGGTCACTTCCAGTCTGGGCTTCAACCAATCTCTGGTTGG 1800
 Db 5245 AAACCTCTCTTACGAGGAGAGGTCACTTCCAGTCTGGGCTTCAACCAATCTCTGGTTGG 5304
 QY 1801 TGCAGCTTCCATGCGAGCGCGAACCAGATGTAGAGATGCTCATCTTCACTCACCAAC 1860
 Db 5305 TGCAGCTTCCATGCGAGCGCGAACCAGATGTAGAGATGCTCATCTTCACTCACCAAC 5364
 QY 1861 CCTCCCATACAGACAGAGCGGCTAAGCGCAGGCTGAGCCAGGGGGTCTCCCGCTTC 1920
 Db 5365 CCTCCCATACAGACAGAGCGGCTAAGCGCAGGCTGAGCCAGGGGGTCTCCCGCTTC 5424
 QY 1921 TTGGCAGCTTCTTACGTTAGCTAGCCAGTTGTCTGGCTTCTCTGAAAGCGCATATAC 1980
 Db 5425 TTGGCAGCTTCTTACGTTAGCTAGCCAGTTGTCTGGCTTCTCTGAAAGCGCATATAC 5484
 QY 1981 CAAATGACTTCCAGAGGCTGACTCATCGAGGCGCAACCTCTGTGTGCGGATAGAGATG 2040
 Db 5485 CAAATGACTTCCAGAGGCTGACTCATCGAGGCGCAACCTCTGTGTGCGGATAGAGATG 5544
 QY 2041 GGC 2043
 Db 5545 GGC 5547

RESULT 8
 US-08-384-616-31
 ; Sequence 31, Application US/08384616
 ; Patent No. 5847101
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKE, Isao


```

QY 1201 AAGAGCTCCACGATGATTAACGAGACTGCTCCAGCCCTGCTCCGGCTGCTGAGCTA 1260
DB 4705 AAAAGCTCCACGATGATTAACGAGACTGCTCCAGCCCTGCTCCGGCTGCTGAGCTA 4764
QY 1261 AAGGATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 4765 AAGGATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4824
QY 1321 AAGCTCTGCTCCGAGATTAACGAGAGTCCCTTTTTCATGCAAGAGTGGGATCAAGAGG 1380
DB 4825 AAGCTCTGCTCCGAGATTAACGAGAGTCCCTTTTTCATGCAAGAGTGGGATCAAGAGG 4884
QY 1381 GTCTGGCGGGAGAGCGGATCATGACAGCCCTGCTCATGTCGAGACAGATCACCGGA 1440
DB 4885 GTCTGGCGGGAGAGCGGATCATGACAGCCCTGCTCATGTCGAGACAGATCACCGGA 4944
QY 1441 CATGTCAAAAACGGTTCCATGAGAGTGTGGGCTTAAGACTGTAGTAAATCATGTGGCAT 1500
DB 4945 CATGTCAAAAACGGTTCCATGAGAGTGTGGGCTTAAGACTGTAGTAAATCATGTGGCAT 5004
QY 1501 GGAACATTCGCCATCAACGATACACAGGAGGCGGCTGACAGGCGCCCTCCAGGCGCAAC 1560
DB 5005 GGAACATTCGCCATCAACGATACACAGGAGGCGGCTGACAGGCGCCCTCCAGGCGCAAC 5064
QY 1561 TATTCAGGAGCGCTGTGGCGGCTGTGGCTGTGAGAGTACGAGAGGTTACGCGGGTGGG 1620
DB 5065 TATTCAGGAGCGCTGTGGCGGCTGTGGCTGTGAGAGTACGAGAGGTTACGCGGGTGGG 5124
QY 1621 GATTTCCACTACGTGACGAGCATGACCATGACAGTAAATGCCCGTCCAGGTTCCA 1680
DB 5125 GATTTCCACTACGTGACGAGCATGACCATGACAGTAAATGCCCGTCCAGGTTCCA 5184
QY 1681 GCGCCCGAATTTCTTCACAGAGTATGGGGTGGGCTGACAGTACGCTCCGGGCTGC 1740
DB 5185 GCGCCCGAATTTCTTCACAGAGTATGGGGTGGGCTGACAGTACGCTCCGGGCTGC 5244
QY 1741 AAACCTCTCTCTACGAGAGAGGTACATTCAGAGTGGGCTCAACCAATACCTGTTGGG 1800
DB 5245 AAGGCTCTCTCTACGAGAGAGGTACATTCAGAGTGGGCTCAACCAATACCTGTTGGG 5304
QY 1801 TCGCAGCTCCCATGCGAGCCCGAACCAGATGTAAGAGTGTCACTTCATGCTCAGCAG 1860
DB 5305 TCGCAGCTCCCATGCGAGCCCGAACCAGATGTAAGAGTGTCACTTCATGCTCAGCAG 5364
QY 1861 CCTCCCAATACAGCAGCAGAGCGGTAACCGGAGCTGGCCAGGGGGTCTCCCGCTCC 1920
DB 5365 CCTCCCAATACAGCAGCAGAGCGGTAACCGGAGTGTGAGGAGGAGGAGTCTCCCGCTCC 5424
QY 1921 TTGGCCAGCTCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
DB 5425 TTGGCCAGCTCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5484
QY 1981 CAAATGACTTCCAGAGCGCTGACCTCATGAGAGCCAACTCTGTGGCGGCAATGAGAT 2040
DB 5485 CAAATGACTTCCAGAGCGCTGACCTCATGAGAGCCAACTCTGTGGCGGCAATGAGAT 5544
QY 2041 GGC 2043
DB 5545 GGC 5547

```

RESULT 9
US-08-904-686A-31

```

; Sequence 31, Application US/08904686A
; Patent No. 598130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKE, Iwao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIYAMA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; CDNA AND ANTIGEN POLYPEPTIDE

```

```

; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
; ADDRESSER: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcleland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 687-0357
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7917 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7862
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..7916
; OTHER INFORMATION: /note="sequence = 1500 - 9416 of
; OTHER INFORMATION: SEQ ID NO: 1"
; US-08-904-686A-31

```

Query Match 87.9%; Score 1796.6; DB 2; Length 7917;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```

QY 1 TGGAGAGGCGCTTTCACAGGCTTCCACGACGATGATGATGATGATGATGATGATGATGAT 60
DB 3505 TGGAGAGGCGCTTTCACAGGCTTCCACGACGATGATGATGATGATGATGATGATGATGAT 3564
QY 61 CAGGAGAGAGCAACTTCCCTCACTGTGTGGCGGTAACAGGCTACTGTGTGGCGGTAAGGCC 120
DB 3565 CAGGAGAGAGCAACTTCCCTCACTGTGTGGCGGTAACAGGCTACTGTGTGGCGGTAAGGCC 3624

```


QY	121	CAGGCCCCACTTCATCATATGGGATTAATATGGAAAGTGTCTTCATACGGCTTAACGCTACT	180
Db	3625	CAGGCCCCACTTCATCATATGGGATTAATATGGAAAGTGTCTTCATACGGCTTAACGCTACT	3684
QY	181	CTGCACGGGGCCAAACACCCTTGCTGTATATGGCTGGGAGCGCTCCAAAAGAGGTCAACCTC	240
Db	3685	CTGCACGGGGCCAAACACCCTTGCTGTATATGGCTGGGAGCGCTCCAAAAGAGGTCAACCTC	3744
QY	241	ACACACCCCATTAACCAAAATTCAATCATGAGCATGATGTCAAGCCGACCTGGAGAGTGTACG	300
Db	3745	ACCCACCCCATTAACCAAAATTCAATCATGAGCATGATGTCAAGCCGACCTGGAGAGTGTACG	3804
QY	301	AGCACTGGGTGCTGTGGTGGGCGGGGTCTTGATAGCTCTGCGCTGATATTTGTTACAA	360
Db	3805	AGCACTGGGTGCTGTGGTGGGCGGAATCTTGAGCTCTGGCGGGATATTTGTTACAA	3864
QY	361	GGCACGCTGTCATTGTGGGTAGATCATCTTGTCTGGGCGGACCGGCTATTTGTTCCGAC	420
Db	3865	GGCACGCTGTCATTGTGGGTAGATCATCTTGTCTGGGCGGACCGGCTATTTGTTCCGAC	3924
QY	421	AGGGAAGTCTCTTACCAAGATTGATGATGGAAGTGGCGGTGCACCTCCCTTAC	480
Db	3925	AGGGAAGTCTCTTACCAAGATTGATGATGGAAGTGGCGGTGCACCTCCCTTAC	3984
QY	481	ATCGAGCAGGGAATGACATCGCGGAGAGTTCAAGCAAAAAGGCTCGGGTTCTGAC	540
Db	3985	ATCGAGCAGGGAATGACATCGCGGAGAGTTCAAGCAAAAAGGCTCGGGTTCTGCA	4044
QY	541	ACAGCCACCAAGCAAGCGGAGGCGCTCTCCGCTCGGTGTGAGTCCAAATGGCGAGCCCTT	600
Db	4045	ACAGCCACCAAGCAAGCGGAGGCTCTCCGCTCGGTGTGAGTCCAAATGGCGAGCCCTT	4104
QY	601	GAGACCTTCTGGGCGAAACATATGTGAACTTCATCAACGGGATTCAGTACTTATAGCAGC	660
Db	4105	GAGACCTTCTGGGCGAAACATATGTGAACTTCATCAACGGGATTCAGTACTTATAGCAGC	4164
QY	661	TTGTCCACATCTGCGGGGAATCCGCGCAFTGTGATCATGATGAGGCTTACAGCCCTGTCT	720
Db	4165	TTATTCACATCTGCGGGGAATCCGCGCAATGATATGATGAGCACTTACAGCCCTTATC	4224
QY	721	ACTAGCCGCTCACCAACCAATCTACCTCTGTCTTAAATCTCTGGGGGAATGGGTAGCC	780
Db	4225	ACCAGCCGCTCACCAACCAATGATCCCTCGTTTAAATCTTGGGGGGTGGGTGGCT	4284
QY	781	GGCCCACTGCTCCCCCAAGTGTCTTCAGCTTTCGTATAGGCGCGGGAATTGCTGTGTGCG	840
Db	4285	GGCCCACTGCTCCCCCAAGCGCCGCTTTCGCTTTCGTGTGGGCGCGGATGCGCGGTGCG	4344
QY	841	GCTGTTTGGCAGATAGGCTTTGGGAAGGTCTTGTGACATCTTGGCGGAGCTATGAGCA	900
Db	4345	GCTGTTTGGCAGATAGGCTTTGGGAAGGTGCTTGTGACATCTTGTGGCGGTTATGAGCA	4404
QY	901	GGAGTGGCAGGCGCGTGTGGCTTTTAAAGTCATGAGCGGGAATGCTCCACCGAG	960
Db	4405	GGAGTGGCGCGGCGCTGTGGCTTTTAAAGTCATGAGCGGGAATGCTCCACCGAG	4464
QY	961	GACCTGTGTTAACTTATCCTCCGTGCAATCCTCTCTCTGATGCCGTGTGTGGGGTGTGTG	1020
Db	4465	GACCTGTGTTAACTTATCCTCTGCAATCCTCTCTCTGATGCCGTGTGTGGGGTGTGTG	4524
QY	1021	TGCGAGCGATATCTGCTGCGACGCTGGGTTCAGGGGAGGGGGCTGTCACTGATGAAC	1080
Db	4525	TGTGAGCAATATCTGCTGTGACATGTGGGTCCGGGAGAGGGGGCTGTGCACTGATGAAC	4584
QY	1081	CGGCTGATAGCGTGTGCGCTCGCGGGGTAAACATGTTTCCCCCGACCTATATGTGACAG	1140
Db	4585	CGGCTGATAGCGTGTGCGCTCGCGGGGTATATGTTTCCCCCGACCTATATGTGCTGAG	4644
QY	1141	AACGAGCGCGCAGCAGTGTCACTAGATCCTCTCGACTTATACCACTATACCACTGTG	1200
Db	4645	AACGAGCGCGCAGCGGTGTACTAGATCCTCTCGACTTATACCACTATACCACTGTGCTG	4704

QY	1201	AAAGAGCCTCCACCAAGTGGATTAAACAGAGAACCTGGCTCCAGCGCCCTGCTCCGGCTCGTGGCTTA	1266
Db	4705	AAAAGGCTCCACCAAGTGGATTAAAGAAAGACTGGCTCCAGCACCCTGTTCCGGCTCGTGGCTTA	4764
QY	1261	AGGATGTTTGGGACTGGATTATGACACAGTTTGGCTGACTTTCAAGACTTGACTTC	1320
Db	4765	AGGATGTTTGGACTGGATTATGACAGGCTGTGTGACTGACTTCAAGAACCTGAGTCCAGTCC	4824
QY	1321	AAGCTCCTGCGCGGATTACCGGGAGTCCCTTTTCTCATGCCAAAGTGGGTACAGGGG	1380
Db	4825	AAAGCTCCGCGCGAGCTACCTGGAGTCCCTTTTCTCGTGCACACCGGGGTACAAAGGA	4884
QY	1381	GTCGCGGGGGGAGACGGCATCATGACACACCGCTCATGTGGAGACAGATCACCGGA	1440
Db	4885	GTCGCGGGGGGAGACGGCATCATCAACCACTGCCATGTGGAGACAGATCACCGGA	4944
QY	1441	CATGTCAAAAAAGTTCCATGAGGATGTTGGGCTTAAGACTGTACTTAACATGTGCAT	1500
Db	4945	CATGTCAAAAAAGTTCCATGAGGATGTGTGGGGCTTAAGACTGTGACCAACATGTGCAT	5004
QY	1501	GGAACTATCCCATCAACGATACACACCGGGCCCTTGCAAGCCCTTCCCAAGGCCAAAC	1560
Db	5005	GGAACTATCCCATCAACGATACACACCGGGCCCTTGCAAGCCCTTCCCAAGGCCCAAC	5064
QY	1561	TATTCGAGGGGCGTGTGGCGGGGTGGCTGAGGAGTACGTGAGAGTTAGCGGGGTGGG	1620
Db	5065	TATTCGAGGGGCGTGTGGCGGGGTGGCTGAGGAGTACGTGAGAGTACGCGGGGTGGG	5124
QY	1621	GATTTCCACTACGTGACGAGCATGACCACTGACCAACGTAATAATGCCCCGTCCAGATTCA	1680
Db	5125	GATTTCCACTACGTGACGAGCATGACCACTGACCAACGTAATAATGCCATCCAGATTCCG	5184
QY	1681	GCCCCCGAATTCCTACAGAAAGTGATGGGGTGCAGCTGACAGGTACGCTCCGGCGTGC	1740
Db	5185	GCTCCTGAATTCCTCGAGAGGTGGACGGAAGTGGCTTGACAGGTACGCTCCGGCGTGC	5244
QY	1741	AAACCTCTCCACCGGAGAGAGGTCAKMTTCCAGGTGGGGCTCAACCAATKCTGGTTGGG	1800
Db	5245	AGGCTCTCCACCGGAGAGAGTTCAKMTTCCAGGTGGGGCTCAACCAATKCTGGTTGGG	5304
QY	1801	TGCGAGCTCCCATGCGAGCCGGAACCGGATGTAGCAGTGTCACTTCCATGCTCACCGAC	1860
Db	5305	TCACAGCTCAATGCGAGCCCGGAACCGGATGTAGCAGTGTCACTTCCATGCTCACCGAC	5364
QY	1861	CCCTCCACATCACAGCAGAGACGGCTTAAGCGAGGCTGGCAGAGGGGTCTCCCTCCCTCC	1920
Db	5365	CCCTCCACATCACAGCAGGAACGGCTTAAGGTTGGCTGGCCAGGGGGTCTCCCTCCCTCC	5424
QY	1921	TTGGCCAGACTCTTCAAGCTAGCCAGTTGTCTGGCCCTTCTCTGAAGGGACATACATTACC	1980
Db	5425	TTGGCCAGACTCTTCAAGCTAGCCAGTTGTCTGGCCCTTCTCTGAAGGGGACATGACTTACC	5484
QY	1981	CAAAATGACTTCCAGAGCGTGAAGCTGATCGAGGCGCAACCTCTGTGGCGGATGAGATG	2040
Db	5485	CACCATGCTCTCCGAGCGTGAAGCTGATCGAGGCGCAACCTCTGTGGCGGAGAGATG	5544
QY	2041	GGC 2043	
Db	5545	GGC 5547	
RESULT 10			
US-09-315-850-31			
; Sequence 31, Application US/09315850			
; Patent No. 6217872			
; GENERAL INFORMATION:			
; APPLICANT: OKAYAMA, Hiroto			
; APPLICANT: FUJIE, Iseo			
; APPLICANT: MORI, Chisato			
; APPLICANT: TAKAMIZAWA, Akahisa			
; APPLICANT: YOSHIDA, Iwao			
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC			
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE			

NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
 ADDRESSEE: Naughton
 STREET: 1725 K St. N.W. Suite 1000
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20006
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/315,850
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/904,686
 FILING DATE: 01-AUG-1997
 APPLICATION NUMBER: US/08/324,977
 FILING DATE: 18-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-167466
 FILING DATE: 25-JUN-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-230921
 FILING DATE: 31-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-305605
 FILING DATE: 09-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/099,706
 FILING DATE: 30-JUL-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/769,996
 FILING DATE: 02-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/635,451
 FILING DATE: 28-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Mcleland, Le-Nhung
 REGISTRATION NUMBER: 31,541
 REFERENCE/DOCKET NUMBER: 9007036
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-0357
 TELEFAX: (202) 659-2930
 INFORMATION FOR SEQ ID NO: 31:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7917 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA from genomic RNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..7862
 NAME/KEY: misc_feature
 LOCATION: 1..7916
 OTHER INFORMATION: /note="sequence = 1500 - 9416 of
 OTHER INFORMATION: SEQ ID NO: 1"
 US-09-315-850-31
 Query Match 87.9%; Score 1796.6; DB 3; Length 7917;
 Best Local Similarity 92.5%; Pred. No. 0;
 Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
 QY 1 TGGGAGGGGCTTTTCAACAGGCTCAACCCAGTGATGCCCACTTCTGTGCCAAAGAAG 60
 DB 3505 TGGGAGAGTGTCTTCAACAGGCTCAACCCATATAGATGCACACTTCTGTGCCAGACCAAG 3564
 QY 61 CAGGAGAGACAACCTTCCCTACCTGATGGGCTACCAAGCTACTGTGTGCGGTAGGCGC 120

DB 3565 CAGGAGAGACAACCTTCCCTACCTGATGATACCAAGCCAGGTGTGCGGCAAGGCT 3624
 QY 121 CAGGCCCACTTCCATCATAGGATCAAAATGTGAATGTCTCATAGGCTTAAAGCTACT 180
 DB 3625 CAGGCCCACTTCCATCATAGGATCAAAATGTGAATGTCTCATAGGCTTAAAGCTACT 3684
 QY 181 CTGCGCGGGGCAACACCTTGTGCTGTATAGGCTGGAGCCGCTCAAAAGAGTCAACCTC 240
 DB 3685 CTGCAAGGGGCCAACACCTTGTGCTGTATAGGCTGGAGCCGCTCAAAAGAGTCAACCTC 3744
 QY 241 ACAACACCCCATTAACCAATTCATCATGATGATGATGATGATGATGATGATGATGATG 300
 DB 3745 ACCACACCCCATTAACCAATTCATCATGATGATGATGATGATGATGATGATGATGATG 3804
 QY 301 AGCACTGGGTGTGTGTGGGCGGGGCTTGTGAGCTGTGGCTGTGGCTGTGGCTGTGGCT 360
 DB 3805 AGCACTGGGTGTGTGTGGGCGGGGCTTGTGAGCTGTGGCTGTGGCTGTGGCTGTGGCT 3864
 QY 361 GGCAGCGTGTATGTGTGGGTAGGATCATCTTGTCCGGGCGGCGCTATTGTTCGCGAC 420
 DB 3865 GGCAGTGTGTATGTGTGGGTAGGATCATCTTGTCCGGGCGGCGCTATTGTTCGCGAC 3924
 QY 421 AGGGAAGTCTCTACCAAGAGTTCGATGATGAAAGATGCGCGTGCACCTCCCTTAC 480
 DB 3925 AGGGAAGTCTCTACCAAGAGTTCGATGATGAAAGATGCGCGTGCACCTCCCTTAC 3984
 QY 481 ATGAGCAGGGAATGCACTCGCCGAGCACTTCAAGCAAAAGCGCTCGGTTGCTGACG 540
 DB 3985 ATGAGCAGGGAATGCACTCGCCGAGCACTTCAAGCAAAAGCGCTCGGTTGCTGACG 4044
 QY 541 ACAGCACAACAACGAGGCGGCGCTGCTCCGAGTGTGAGTGTCCAGTGGGAGGCGCT 600
 DB 4045 ACAGCACAACAACGAGGCGGCGCTGCTCCGAGTGTGAGTGTCCAGTGGGAGGCGCT 4104
 QY 601 GAGACCTTGTGGGCGGGAACAATGTGAATCTTATCAGCGGATACAGTACTTATGAGAGC 660
 DB 4105 GAGACCTTGTGGGCGGGAACAATGTGAATCTTATCAGCGGATACAGTACTTATGAGAGC 4164
 QY 661 TTGTCCACTGTGCTGTGGGAATCCCGGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 4165 TTATCCACTGTGCTGTGGGAATCCCGGATGATGATGATGATGATGATGATGATGATGAT 4224
 QY 721 ACTAGCCGCTACCAACCAATCTAACCCTGCTTAATCTGTGGGAGATGGGATAGGC 780
 DB 4225 ACTAGCCGCTACCAACCAATCTAACCCTGCTTAATCTGTGGGAGATGGGATAGGC 4284
 QY 781 GCCCAACTGCTCCCGCAGTGTGCTGCTTGTGATGAGGCGCGCATTTGCTGGTGGC 840
 DB 4285 GCCCAACTGCTCCCGCAGTGTGCTGCTTGTGATGAGGCGCGCATTTGCTGGTGGC 4344
 QY 841 GCTGTGTGAGCATAGGCTTGGGAAGTGTCTTGTGATGATGATGATGATGATGATGATGAT 900
 DB 4345 GCTGTGTGAGCATAGGCTTGGGAAGTGTCTTGTGATGATGATGATGATGATGATGATGAT 4404
 QY 901 GGAATGCAAGGCGGCTGTGCTTGAAGTCAATGATGATGATGATGATGATGATGATGATG 960
 DB 4405 GGAATGCAAGGCGGCTGTGCTTGAAGTCAATGATGATGATGATGATGATGATGATGATG 4464
 QY 961 GACCTGTAACTTACTCCCTGCAATCTCTCTGCTGTGAGGCTGTGAGGCTGTGAGGCT 1020
 DB 4465 GACCTGTAACTTACTCCCTGCAATCTCTCTGCTGTGAGGCTGTGAGGCTGTGAGGCT 4524
 QY 1081 CGGCTGATAGCTTGTGCTGTGGGAGTAACTATGTTTCCCAACCACTATGTCAGAG 1140
 DB 4585 CGGCTGATAGCTTGTGCTGTGGGAGTAACTATGTTTCCCAACCACTATGTCAGAG 4644
 QY 1141 AGCGAGCGGAGAGTGTCACTAGATCTCTCCGACCTTACTATCAACCACTGTG 1200

```

Db      4645 AGCGAGCCGCGAGCGCTGTGTACTCAGACRCTCTCCAGCCTTACCATCATCTCAGCTGCTG 4704
Qy      1201 AAGAGGCTCCACGAGTGATTAAACGAGAGCTGCTCCAGCCCTGCTCCGCTCTGAGGTA 1260
Db      4705 AAAAGGCTCCACGAGTGATTAAAGAGAGCTGCTCCAGCCCTGCTCCGCTCTGAGGTA 4764
Qy      1261 AGGAGTGTGGGAGTGAATGACACAGTTTGGCTGACTTCAAGACCTGGCTCCAGTTC 1320
Db      4765 AGGAGTGTGGGAGTGAATGACACAGTTTGGCTGACTTCAAGACCTGGCTCCAGTTC 4824
Qy      1321 AAGCTCCGCGCGAGTTACCGCGGAGTCCCTTTTCTCATGTCACAGTGGGAGCAAGGGG 1380
Db      4825 AAGCTCCGCGCGAGTTACCGGAGTCCCTTTTCTCATGTCACAGGCGGAGTCAAGGGA 4884
Qy      1381 GTCTGGCGGGGAGAGCGGATCATGAGACCACTGCTCATGTGAGCAAGATCAACCGGA 1440
Db      4885 GTCTGGCGGGGAGAGCGGATCATGAGACCACTGCTCATGTGAGCAAGATCAACCGGA 4944
Qy      1441 CATGTCAAAAACGGTTTCATGAGATCGTTGGGCTTAAGAAGCTGTATGTAACATGTGGCAT 1500
Db      4945 CATGTCAAAAACGGTTTCATGAGATCGTTGGGCTTAAGAAGCTGTATGTAACATGTGGCAT 5004
Qy      1501 GGAACATTTCCCATGAAGCATACACACAGGCGCCCTGACAGCCCTCCCGACGCGCAAC 1560
Db      5005 GGAACATTTCCCATGAAGCATACACACAGGCGCCCTGACAGCCCTCCCGACGCGCAAC 5064
Qy      1561 TATTCAGAGGCGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGAGTTACGCGGGTGGGG 1620
Db      5065 TATTCAGAGGCGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGAGTTACGCGGGTGGGG 5124
Qy      1621 GATTTCCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCCCGTGCAGGTTCCA 1680
Db      5125 GATTTCCACTACGTGACGAGCATGACCACTGACAAAGTAAATGCCCGTGCAGGTTCCG 5184
Qy      1681 GCCCGCGAATTTCTTCAAGAAAGTGAATGGGGTGGCGGCTGACAGGATGCGTCCGGCGTGC 1740
Db      5185 GCTCTGAATTTCTTCTCGAGAGTGAACGAGTGGCGGTTGACAGAGTACGCTCCGGCGTGC 5244
Qy      1741 AAACCTCTCTACGAGGAGAGGTCAATTCACAGTCCGAGCTCAACCAATCTGTTGGG 1800
Db      5245 AGGCTCTCTCTACGAGGAGAGGTCAATTCACAGTCCGAGCTCAACCAATCTGTTGGG 5304
Qy      1801 TCGAGCTCCCATGCGAGCCGACCCGATGTAGACATGCTCACTTCCATGCTCACCGAC 1860
Db      5305 TCGAGCTCCCATGCGAGCCGACCCGATGTAGACATGCTCACTTCCATGCTCACCGAC 5364
Qy      1861 CCCTCCCATGACAGACGAGACGCTTAAGCGGAGCGGCTGCGACAGGCGGCTCCCGCTCC 1920
Db      5365 CCCTCCCATGACAGACGAGACGCTTAAGCGGAGTGGCCAGAGGCGGCTCCCGCTCC 5424
Qy      1921 TTGGCAGGCTTTCAGCTAGCAGTGTGTGCGCCTTCTCGAAGGCGCATATTAAC 1980
Db      5425 TTGGCAGGCTTTCAGCTAGCAGTGTGTGCGCCTTCTCGAAGGCGCATATTAAC 5484
Qy      1981 CAAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTTCCTGTGCGGCGATGAGATG 2040
Db      5485 CACCATGTCTCTCCGAGCGCTGACCTCATGAGGCCAACCTTCCTGTGCGGCGAGAGATG 5544
Qy      2041 GGC 2043
Db      5545 GGC 5547

```

```

RESUT 11
US-08-324-977-13
; Sequence 13, Application US/08324977
; Patent No. 5747339

```

```

; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKE, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao

```

```

; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McJeland &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9030 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA from genomic RNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..9030
; OTHER INFORMATION: /note="sequence = 333 - 9362 of
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..9030
; US-08-324-977-13

```

```

Query Match      87.9%; Score 1796.6; DB 1; Length 9030;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```

```

Qy      1  TGGAGGGCGCTTTCACAGGCTTCAACGATGATGCCCACTTCTGTGCTCAACAAG 60
Db      4672 TGGAGAGTGTCTTCACAGGCTTCAACGATGATGACACTTCTGTGCTCAACAAG 4731
Qy      61  CAGGAGAGACAATTTCCCTCACTGAGGCGTACAGGCTACTGTGTGCGCTAGGGCC 120
Db      4732 CAGGAGAGACAATTTCCCTCACTGAGGCGTACATCCAAAGCCAGGTGTGCGCTAGGGCT 4791

```

QY 121 CAGGCCCACTTCATCATGAGATCAATGTGAAAGTGTCTCATACGCTAAAGCTTACT 180
 DB 4792 CAGGCCCACTTCATCATGAGATCAATGTGAAAGTGTCTCATACGCTAAAGCTTACT 4851
 QY 181 CTGCGCGGGCCCAACACCTTGTCTGTATAGGCTGGAGCCGCTCCAAAACAGGTCACTTC 240
 DB 4852 CTGCGCGGGCCCAACACCTTGTCTGTATAGGCTGGAGCCGCTCCAAAACAGGTCACTTC 4911
 QY 241 ACACACCCCATTAACAAATTCATCATGAGATCAATGTGAAAGTGTCTCATACGCTAAAGCTTACT 300
 DB 4912 ACACACCCCATTAACAAATTCATCATGAGATCAATGTGAAAGTGTCTCATACGCTAAAGCTTACT 4971
 QY 301 AGCACTGGGTGCTGTGTGGGCGGGGCTCTTGACAGCTGTGGCTGCTGCTGCAACA 360
 DB 4972 AGCACTGGGTGCTGTGTGGGCGGGGCTCTTGACAGCTGTGGCTGCTGCTGCAACA 5031
 QY 361 GGCAGCGGTGCTATGTGTGGATCATCTTGTCCGGGCGGCTATGTGTCCGAC 420
 DB 5032 GGCAGCGGTGCTATGTGTGGATCATCTTGTCCGGGCGGCTATGTGTCCGAC 5091
 QY 421 AGGGAAGCTTCTTACAGAGATTCATGAGATGGAAGATGGCGCTGCGACCTTCCCTTAC 480
 DB 5092 AGGGAAGCTTCTTACAGAGATTCATGAGATGGAAGATGGCGCTGCGACCTTCCCTTAC 5151
 QY 481 ATCGACAGGGAATGACAGCTCGCGAGCAGTTCAAGCAAAAAGCGCTGGGTGCTGAC 540
 DB 5152 ATCGACAGGGAATGACAGCTCGCGAGCAGTTCAAGCAAAAAGCGCTGGGTGCTGAC 5211
 QY 541 ACAGCCCAACCAAGCGAGGCGCTGCTCCGTGTGTGAGTCCAAATGGCGAGCCCTT 600
 DB 5212 ACAGCCCAACCAAGCGAGGCGCTGCTCCGTGTGTGAGTCCAAATGGCGAGCCCTT 5271
 QY 601 GAGACCTTCTGGGCGAAAACATGTGGAATTCATCAAGCGGATACAGTACTTAAAGGAC 660
 DB 5272 GAGACCTTCTGGGCGAAAACATGTGGAATTCATCAAGCGGATACAGTACTTAAAGGAC 5331
 QY 661 TTGTCACTCTGCTGGGAATCCCGGATGTGATCACTGATGAGCTTCAAGCTCTGTC 720
 DB 5332 TTGTCACTCTGCTGGGAATCCCGGATGTGATCACTGATGAGCTTCAAGCTCTGTC 5391
 QY 721 ACTAGCCCGCTCACCAACCAATCTACCTCTGCTTAAATCATCTGGGGGATGGGTAGCC 780
 DB 5392 ACTAGCCCGCTCACCAACCAATCTACCTCTGCTTAAATCATCTGGGGGATGGGTAGCC 5451
 QY 781 GCCCACTGCTCTCCCGAGTCTGCTTCACTGATGAGCGCGGCACTTGTGTGTG 840
 DB 5452 GCCCACTGCTCTCCCGAGTCTGCTTCACTGATGAGCGCGGCACTTGTGTGTG 5511
 QY 841 GCTGTGGCAGATAGGCGCTTGGGAAGGTGTGGAATCTTGGCGGCTTATGAGCA 900
 DB 5512 GCTGTGGCAGATAGGCGCTTGGGAAGGTGTGGAATCTTGGCGGCTTATGAGCA 5571
 QY 901 GAGATGAGGCGCGCTGCTGCTTAAAGTCAATGAGCGGGAATGCTTCAACCGAG 960
 DB 5572 GAGATGAGGCGCGCTGCTGCTTAAAGTCAATGAGCGGGAATGCTTCAACCGAG 5631
 QY 961 GACCTGTAACTTACTCTCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 DB 5632 GACCTGTAACTTACTCTCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 5691
 QY 1021 TGCGCAGGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 DB 5692 TGCGCAGGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5751
 QY 1081 CGGCTGATAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 DB 5752 CGGCTGATAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5811
 QY 1141 AGCGACCGCGCAGCAGTGTCTCATGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 DB 5812 AGCGACCGCGCAGCAGTGTCTCATGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 5871

QY 1201 AAGAGCTTCACCAAGTGAATTAACAGAGCTGCTCCACGCGCTGCTCGGCTGCTGCTGCTA 1260
 DB 5872 AAGAGCTTCACCAAGTGAATTAACAGAGCTGCTCCACGCGCTGCTCGGCTGCTGCTGCTA 5931
 QY 1261 AAGAGTGTGGGACTGATATGACAGTGTGGCTGACTTCAAGACCTGCTCCAGTCC 1320
 DB 5932 AAGAGTGTGGGACTGATATGACAGTGTGGCTGACTTCAAGACCTGCTCCAGTCC 5991
 QY 1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGTAAAGGGG 1380
 DB 5992 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGCGCAAGTGGTAAAGGGG 6051
 QY 1381 GTCTGCGCGGAGAGCGGATTCATGAGACCACTGCTCATGTGAGCAGATCAACCGGA 1440
 DB 6052 GTCTGCGCGGAGAGCGGATTCATGAGACCACTGCTCATGTGAGCAGATCAACCGGA 6111
 QY 1441 CATGTCAAAAACGTTCCATGAGATCGTTGGGCTTAAAGCTGTATGAATGTGGCAT 1500
 DB 6112 CATGTCAAAAACGTTCCATGAGATCGTTGGGCTTAAAGCTGTATGAATGTGGCAT 6171
 QY 1501 GGAACATTCCTCATGAGCAGATCAACCAAGGCGCTGACGCGCTCCACGCGCAAC 1560
 DB 6172 GGAACATTCCTCATGAGCAGATCAACCAAGGCGCTGACGCGCTCCACGCGCAAC 6231
 QY 1561 TATTCAGGCGCTGCTGCGGCTGCTGCTGAGAGTACGTGAGGTTACGCGGCTGAG 1620
 DB 6232 TATTCAGGCGCTGCTGCGGCTGCTGCTGAGAGTACGTGAGGTTACGCGGCTGAG 6291
 QY 1621 GATTTTCATCATGAGCAGATCAACCTGACCAAGTAAATGCCCGTTCAGGTTCCA 1680
 DB 6292 GATTTTCATCATGAGCAGATCAACCTGACCAAGTAAATGCCCGTTCAGGTTCCA 6351
 QY 1681 GCGCGCGAATCTTCAAGAGTGAATGGGCTGCGGCTGACAGGATGCTCGGCGTGC 1740
 DB 6352 GCGCGCGAATCTTCAAGAGTGAATGGGCTGCGGCTGACAGGATGCTCGGCGTGC 6411
 QY 1741 AAACCTCTCTTACGAGGAGGCTCATATTCAGTCTGCGGCTTCAACCAATATCTGTTGG 1800
 DB 6412 AAACCTCTCTTACGAGGAGGCTCATATTCAGTCTGCGGCTTCAACCAATATCTGTTGG 6471
 QY 1801 TCGAGCTTCCATGCGAGCGCGCAACCGGATGTGAGATGCTCATCTTCAATGCTCACCGAC 1860
 DB 6472 TCGAGCTTCCATGCGAGCGCGCAACCGGATGTGAGATGCTCATCTTCAATGCTCACCGAC 6531
 QY 1861 CCTTCCCATATACAGCAGAGCGGCTTAAGCCGCAAGCTGCGAGGCGCTTCCCTCC 1920
 DB 6532 CCTTCCCATATACAGCAGAGCGGCTTAAGCCGCAAGCTGCGAGGCGCTTCCCTCC 6591
 QY 1921 TTGGCAGCTTCTTACGCTAGCAGTGTGCTGCGCTTCTCGAAGGCGCATATTAAC 1980
 DB 6592 TTGGCAGCTTCTTACGCTAGCAGTGTGCTGCGCTTCTCGAAGGCGCATATTAAC 6651
 QY 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCAACCTTCTGCGGCGATGAGATG 2040
 DB 6652 CAAATGACTTCCAGACGCTGACCTCATGAGGCAACCTTCTGCGGCGATGAGATG 6711
 QY 2041 GGC 2043
 DB 6712 GGC 6714

RESULT 12
 US-08-384-616-13
 ; Sequence 13, Application US/08384616
 ; Patent No. 5847101
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKE, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE

```

1      NUMBER OF SEQUENCES: 50
2      CORRESPONDENCE ADDRESS:
3      ADDRESSEE: Armstrong, Westernman, Hattori, McIeland
4      ADDRESSEE: Naughton
5      STREET: 1725 K St. N.W. Suite 1000
6      CITY: Washington
7      STATE: D.C.
8      COUNTRY: U.S.A.
9      ZIP: 20006
10     COMPUTER READABLE FORM:
11     MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
12     COMPUTER: IBM PC compatible
13     OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
14     SOFTWARE: ASCII
15     CURRENT APPLICATION DATA:
16     APPLICATION NUMBER: US/08/384,616
17     FILING DATE:
18     CLASSIFICATION: 424
19     PRIOR APPLICATION DATA:
20     APPLICATION NUMBER: US 07/769,996
21     FILING DATE: 02-OCT-1991
22     APPLICATION NUMBER: JP 2-167466
23     FILING DATE: 25-JUN-1990
24     PRIOR APPLICATION DATA:
25     APPLICATION NUMBER: JP 2-230921
26     FILING DATE: 31-AUG-1990
27     PRIOR APPLICATION DATA:
28     APPLICATION NUMBER: JP 2-305605
29     FILING DATE: 09-NOV-1990
30     PRIOR APPLICATION DATA:
31     APPLICATION NUMBER: US 07/635,451
32     FILING DATE: 28-DEC-1990
33     ATTORNEY/AGENT INFORMATION:
34     NAME: Stevens-Smith, Theresa M.
35     REGISTRATION NUMBER: 36,281
36     REFERENCE/DOCKET NUMBER: 900703B
37     TELECOMMUNICATION INFORMATION:
38     TELEPHONE: (202) 659-2930
39     TELEFAX: (202) 867-0357
40     TELEX: 440142
41     INFORMATION FOR SEQ ID NO: 13:
42     SEQUENCE CHARACTERISTICS:
43     LENGTH: 9030 base pairs
44     TYPE: nucleic acid
45     STRANDEDNESS: single
46     TOPOLOGY: linear
47     MOLECULE TYPE: cDNA from genomic RNA
48     FEATURE:
49     NAME/KEY: misc feature
50     LOCATION: 1..9030
51     OTHER INFORMATION: /note: "sequence = 333 - 9362 of
52     OTHER INFORMATION: SEQ ID NO: 1"
53     NAME/KEY: CDS
54     LOCATION: 1..9030
55     US-08-384-616-13

```

```

Db 5932 AGGATGTTGGAGTGAATGACAGCGGTGTGACTGACTTCAAGACCTGGCTCCAGTCC 5991
Qy 1321 AAGCTCCGCGCGCATTAACCGCGAGTCCCTTTTCTCATGCAACGAGGATCAAGGGG 1380
Db 5992 AAGCTCTGCGCGAGTACTGAGAGTCCCTTTTCTGTCGCAACGCGGATCAAGGGG 6051
Qy 1381 GTCTGGCGGGAGAGCGGATCATGAGAGCAACCTGCTCATGTGAGCAAGATCAACCGA 1440
Db 6052 GTCTGGCGGGAGAGCGGATCATGCAACCACTGCTCATGTGAGCAAGATCAACCGA 6111
Qy 1441 CATGTCAAAAAGGTTTCATGAGATCGTTGGGCTTAAGACCTGTATGTAACATGTGGCAT 1500
Db 6112 CATGTCAAAAAGGTTTCATGAGATCGTGGGCTTAAGACCTGTGAGCAACGATGGCAT 6171
Qy 1501 GGAACATTCCTCCATTAAGCATACACACGCGGCTTGAAGCGCTCCCGACGCGCAAC 1560
Db 6172 GGAACATTCCTCCATTAAGCATACACACGCGGCTTGAAGCGCTCCCGACGCGCAAC 6231
Qy 1561 TATTCAGGGGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGAGTTACGCGGGTGGG 1620
Db 6232 TATTCAGGGGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGAGTTACGCGGGTGGG 6291
Qy 1621 GATTTCACTACGTACGAGCATGACCACTGACAAAGTAAAGTGCCTGCGAGGTTCCG 1680
Db 6292 GATTTCACTACGTACGAGCATGACCACTGACAAAGTAAAGTGCCTGCGAGGTTCCG 6351
Qy 1681 GCCCGCAATTTTTCACAGAAATGAGTGGGGTGGCTGACAGATGAGCTCCGCGTGC 1740
Db 6352 GCTCTGAATTTCTTCGAGAGTGAACGAGTGGCGGTTGCAAGGATGACCTCCGCGTGC 6411
Qy 1741 AAACCTCTCTACGAGAGAGGTCAATTCAGGTGCGGCTCAACCAATACCTGGTGGG 1800
Db 6412 AGGCTCTCTCTACGAGAGAGGTCAATTCAGGTGCGGCTCAACCAATACCTGGTGGG 6471
Qy 1801 TCGAGCTCCCATGCGAGCGCCGACCGAGTGTAGAGTGTCACTTCCATGCTCACGAC 1860
Db 6472 TCGAGCTCCCATGCGAGCGCCGACCGAGTGTAGAGTGTCACTTCCATGCTCACGAC 6531
Qy 1861 CCTCTCCCATGACGACGAGAGCGGTAAAGCGGATGCGGCGAGGGGTCTCCCCCTCC 1920
Db 6532 CCTCTCCCATGACGACGAGAGCGGTAAAGCGGTAAAGCGGATGCGGCGAGGGGTCTCCCCCTCC 6591
Qy 1921 TTGGCACTCTTTCAGTACGACGATGTCTGCGCTTCTCGAAGCGCATTAATACC 1980
Db 6592 TTGGCACTCTTTCAGTACGACGATGTCTGCGCTTCTCGAAGCGCATTAATACC 6651
Qy 1981 CAAATGACTTCCGACGAGTGTCACTCATGAGCGCAACCTCTGTGGCGGATGAGATG 2040
Db 6652 CACCATGTCTCTCCGAGCGCTGACCTCATGAGCGCAACCTCTGTGTGGCGGAGATG 6711
Qy 2041 GGC 2043
Db 6712 GGC 6714

```

```

RESULT 13
US-08-904-686A-13
Sequence 13, Application US/08904686A
Patent No. 5998130
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hioto
APPLICANT: FUKE, Iwao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSES: Armstrong, Westerman, Hattori, Mclelland &
ADDRESSES: Naughton
STREET: 1725 K St. N.W. Suite 1000

```

```

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Discrete, 3.5 in, 1.44Mb
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 9030 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..9030
OTHER INFORMATION: /note: "sequence = 333 - 9362 of
OTHER INFORMATION: SEQ ID NO: 1"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..9030
US-08-904-686A-13
Query Match 87.9%; Score 1796.6; DB 2; Length 9030;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```

QY 181 CTGCGGGGCAACACCTTGCTGTATAGGCTGGAGCCGCTCAAAACGAGTCAACCTC 240
 Db 4852 CTGACGGGGCCCAACACCTTGCTGTATAGGCTGGAGCCGCTCAAAACGAGTCAACCTC 4911
 QY 241 ACAACCCCCATPACCAATTCATCATGGCATGCAAGTCAAGCTGAGGTGCTCAG 300
 Db 4912 ACCCAACCCCATPACCAATTCATCATGGCATGCAAGTCAAGCTGAGGTGCTCAG 4971
 QY 301 AGCACTGGGATGCTGGTGGGCGGGGCTTGAGCTTGAGCTGAGCTGATTTGCTTGAACA 360
 Db 4972 AGCACTGGGATGCTGGTGGGCGGGGCTTGAGCTTGAGCTGAGCTGATTTGCTTGAACA 5031
 QY 361 GGCAGCGTGTCTTACAGAGATTCATCTTGTCCGGCGGCGCTATTTGTTCCGAC 420
 Db 5032 GGCAGTGTGTCTTACAGAGATTCATCTTGTCCGGCGGCGCTATTTGTTCCGAC 5091
 QY 421 AGGGAAGTCTTACAGAGATTCATCTTGTCCGGCGGCGCTATTTGTTCCGAC 480
 Db 5092 AGGGAAGTCTTACAGAGATTCATCTTGTCCGGCGGCGCTATTTGTTCCGAC 5151
 QY 481 ATCGACAGGGAATGAGCTGCGCGAGAGTTCAGACAAAACGCTCGGTTGCTGAG 540
 Db 5152 ATCGACAGGGAATGAGCTGCGCGAGAGTTCAGACAAAACGCTCGGTTGCTGAG 5211
 QY 541 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCGTTGAGAGTCCAGTGGCGAGCCCT 600
 Db 5212 ACAGCCACCAAGCAAGCGAGGCGCTGCTCCGTTGAGAGTCCAGTGGCGAGCCCT 5271
 QY 601 GAGACCTTCTGGGCGAAACACATGTGGAATTCATCAGCGGGAATCAGTACTTACAGAG 660
 Db 5272 GAGACCTTCTGGGCGAAACACATGTGGAATTCATCAGCGGGAATCAGTACTTACAGAG 5331
 QY 661 TTGTCACCTGCTGGGGAATCCCGGATGTGATCAGTATGAGCTGAGCTCTGTC 720
 Db 5332 TTATATCATCTGCTGGGGAATCCCGGATGTGATCAGTATGAGCTGAGCTCTGTC 5391
 QY 721 ACTAGCCGCTGACCAACCAATCTACCTCTGCTTAAATCATCTGGGGGATGGGTAGCC 780
 Db 5392 ACTAGCCGCTGACCAACCAATCTACCTCTGCTTAAATCATCTGGGGGATGGGTAGCC 5451
 QY 781 GCCCACTGCTCCCCAGTGTCTTCAAGCTTCTGTAAGGCGCGGCACTTCTGTGTG 840
 Db 5452 GCCCACTGCTCCCCAGTGTCTTCAAGCTTCTGTAAGGCGCGGCACTTCTGTGTG 5511
 QY 841 GCTGTTGGAGCATGAGGCTTGGAAAGGCTTGTGAGCATCTTGGCGGCTATGAGCA 900
 Db 5512 GCTGTTGGAGCATGAGGCTTGGAAAGGCTTGTGAGCATCTTGGCGGCTATGAGCA 5571
 QY 901 GGAAGTGGAGGCGGCTGCTGAGCTTGTAAAGTCAATGAGCGGCAATGCGCTCAACGAG 960
 Db 5572 GGAAGTGGAGGCGGCTGCTGAGCTTGTAAAGTCAATGAGCGGCAATGCGCTCAACGAG 5631
 QY 961 GACCTGCTTAACTTACTCTCTGCAATCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 5632 GACCTGCTTAACTTACTCTCTGCAATCTCTCTGCTGCTGCTGCTGCTGCTGCTG 5691
 QY 1021 TGCGAGCGATCTGCTGCTGAGAGGCTTCCAGGGAGGGGCTGTGAGTGAATGAAC 1080
 Db 5692 TGCGAGCGATCTGCTGCTGAGAGGCTTCCAGGGAGGGGCTGTGAGTGAATGAAC 5751
 QY 1081 CGGCTGATPAGCTTGGCTGCGGGGTAACATGTTTCCCGCAGCACTATGTGCAGAG 1140
 Db 5752 CGGCTGATPAGCTTGGCTGCGGGGTAACATGTTTCCCGCAGCACTATGTGCAGAG 5811
 QY 1141 AGCGAGCGCGAGCACTGTCTCACTGATCTCTCTGCTGCTTACTATCACTCACTGCTG 1200
 Db 5812 AGCGAGCGCGAGCACTGTCTCACTGATCTCTCTGCTGCTTACTATCACTCACTGCTG 5871
 QY 1201 AAGAGGCTTCCAGGATTAACGAGCTGCTCCAGCGCGCTGCTGCTGCTGCTGCTG 1260
 Db 5872 AAGAGGCTTCCAGGATTAACGAGCTGCTCCAGCGCGCTGCTGCTGCTGCTGCTG 5931
 QY 1261 AGGAGTGTGGAGCTGATATGACAAGTGTGGCTGACTTCAAGACTGCTCCAGTCC 1320

Db 5932 AGGAGTGTGGAGCTGATATGACAAGTGTGACTGACTTCAAGACTGCTCCAGTCC 5991
 QY 1321 AAGCTCTGCGCGATPACCGGAGTCCCTTTTCTCATGCCAACGTTAGGTAACAAGGG 1380
 Db 5992 AAGCTCTGCGCGAGTACCTGAGTCCCTTTTCTCATGCCAACGTTAGGTAACAAGGG 6051
 QY 1381 GTTGGCGGGAGAGCGGATCATGACAGCACTGCTCATGTGAGCAAGATCACCGGA 1440
 Db 6052 GTTGGCGGGAGAGCGGATCATGACAGCACTGCTCATGTGAGCAAGATCACCGGA 6111
 QY 1441 CATGCAAAAACGTTCCATGAGATGCTTGGGCTTAAGACTGTGTAATGATGGAT 1500
 Db 6112 CATGCAAAAACGTTCCATGAGATGCTTGGGCTTAAGACTGTGTAATGATGGAT 6171
 QY 1501 GGAACATTCCTCCATCAACGATACCAACGAGGCGCTGCAAGCTCCCGACGCAAC 1560
 Db 6172 GGAACATTCCTCCATCAACGATACCAACGAGGCGCTGCAAGCTCCCGACGCAAC 6231
 QY 1561 TATTCAGGGGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGGTTACCGGGTGGG 1620
 Db 6232 TATTCAGGGGCTGTGGCGGGTGGCTGCTGAGAGTACGTGAGGTTACCGGGTGGG 6291
 QY 1621 GATTTCCACTACGTGACGAGATGACCACTGACCAACGTAATGCTCCGTGCTGCTCA 1680
 Db 6292 GATTTCCACTACGTGACGAGATGACCACTGACCAACGTAATGCTCCGTGCTGCTCA 6351
 QY 1681 GCGCCCGAATCTTCAAGAGTGAATGGGGTGGCGCTGCAAGATGCTCCGCGTGC 1740
 Db 6352 GCTCTGAATCTTCTGAGAGTGAATGGGGTGGCGCTGCAAGATGCTCCGCGTGC 6411
 QY 1741 AAACCTCTCTACGGGAGAGTCAATTCAGGTCGGGCTCAACCAATACCTGTTGGG 1800
 Db 6412 AAGCTCTCTACGGGAGAGTCAATTCAGGTCGGGCTCAACCAATACCTGTTGGG 6471
 QY 1801 TCGAGCTTCCATGCGAGCTCCGAACCGGATGTAGAGTGTCACTTCCATGCTACCGAC 1860
 Db 6472 TCGAGCTTCCATGCGAGCTCCGAACCGGATGTAGAGTGTCACTTCCATGCTACCGAC 6531
 QY 1861 CCTTCCCATCAACGAGAGCGGCTAAGCGAGGCTGAGCGAGGGGCTTCCCGCTCC 1920
 Db 6532 CCTTCCCATCAACGAGAGCGGCTAAGCGAGGCTGAGCGAGGGGCTTCCCGCTCC 6591
 QY 1921 TTGGCAGCTCTTACGCTAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
 Db 6592 TTGGCAGCTCTTACGCTAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6651
 QY 1981 CAAATGACTTCCAGAGCTGACTTCAATGAGGCAACCTCTGTGCGGCTATGATG 2040
 Db 6652 CAAATGACTTCCAGAGCTGACTTCAATGAGGCAACCTCTGTGCGGCTATGATG 6711
 QY 2041 GGC 2043
 Db 6712 GGC 6714

RESULT 14
 US-09-315-850-13
 ; Sequence 13, Application US/09315850
 ; Patent No. 6217872
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKE, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
 ; ADDRESS: Naughton
 ; STREET: 1725 K St., N.W., Suite 1000


```

Qy 1261 AGGATGTTGGGAGTGAATATGCAAGTTTGGCTGACTTCAAGACTGCTCCAGTCC 1320
Db 5932 AGGATGTTGGGAGTGAATATGCAAGTTTGGCTGACTTCAAGACTGCTCCAGTCC 5991
Qy 1321 AAGCTCCGCGCGGAGTACCGGGAGTCCCTTTTCTCATGCGCAAGTGGGTAACAAGGG 1380
Db 5992 AAGCTCCGCGCGGAGTACCGGGAGTCCCTTTTCTCATGCGCAAGTGGGTAACAAGGG 6051
Qy 1381 GTCTGGCGGGAGAGCGGATCATGAGAGCAACCTGCTCATGTGAGAGCAAGATCAACGGA 1440
Db 6052 GTCTGGCGGGAGAGCGGATCATGAGAGCAACCTGCTCATGTGAGAGCAAGATCAACGGA 6111
Qy 1441 CATGTCAAAAACGTTTCATGAGATCGTTGGGCTTAAAGACTGTAGTAACTGTGGCAT 1500
Db 6112 CATGTCAAAAACGTTTCATGAGATCGTTGGGCTTAAAGACTGTAGTAACTGTGGCAT 6171
Qy 1501 GGAATATTTCCCATCAACGATACCAACGAGGAGGCTTGAAGGCTTCCCGGAGGCAAC 1560
Db 6172 GGAATATTTCCCATCAACGATACCAACGAGGAGGCTTGAAGGCTTCCCGGAGGCAAC 6231
Qy 1561 TATTCAGGCGGCTGTGGCGGGTGGCTGTGAGAGTACGTGAGAGTTACGGGGTGGG 1620
Db 6232 TATTCAGGCGGCTGTGGCGGGTGGCTGTGAGAGTACGTGAGAGTTACGGGGTGGG 6291
Qy 1621 GATTTCCACTAGTGAAGCATGACCACTGACCAAGTAAATGCCCGTGCAGGTTCCA 1680
Db 6292 GATTTCCACTAGTGAAGCATGACCACTGACCAAGTAAATGCCCGTGCAGGTTCCA 6351
Qy 1681 GCGCCCGAATTTCTTCAAGAAAGTGAAGTGGGGTGGCTTCAAGAGTACGCTCCGGCTGC 1740
Db 6352 GCTCTGAAATTTCTTCAAGAAAGTGAAGTGGGGTGGCTTCAAGAGTACGCTCCGGCTGC 6411
Qy 1741 AAACCTCTCTCAAGGAGGAGGTCACATTCAGAGTGGGCTCAACCAATATCTGTTGG 1800
Db 6412 AGGCTCTCTCAAGGAGGAGGTCACATTCAGAGTGGGCTCAACCAATATCTGTTGG 6471
Qy 1801 TCGAGCTCCCATGCGAGGCCCAACCGGATGTAGAGTGTCTCACTTCCATGCTCAACG 1860
Db 6472 TCGAGCTCCCATGCGAGGCCCAACCGGATGTAGAGTGTCTCACTTCCATGCTCAACG 6531
Qy 1861 CCTTCCCATCATCAAGAGAGCGGTTAAGCGGAGGCTGGCCAGGGGGTCTCCCTCTCC 1920
Db 6532 CCTTCCCATCATCAAGAGAGCGGTTAAGCGGAGGCTGGCCAGGGGGTCTCCCTCTCC 6591
Qy 1921 TTGGCAGCTTTCAGCTAGCTAGTGTGTGGGCTTCTTCAAGGCGCATACATTACC 1980
Db 6592 TTGGCAGCTTTCAGCTAGCTAGTGTGTGGGCTTCTTCAAGGCGCATACATTACC 6651
Qy 1981 CAAATGACTTCCAGAGCGCTGACCTCATGAGGCCAACCTCTGTGGCGGATGATG 2040
Db 6652 CAAATGACTTCCAGAGCGCTGACCTCATGAGGCCAACCTCTGTGGCGGATGATG 6711
Qy 2041 GGC 2043
Db 6712 GGC 6714

```

```

ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333..9362
US-08-324-977-1

Query Match      87.9%; Score 1796.6; DB 1; Length 9416;
Best Local Similarity 92.5%; Pred. No. 0; Mismatches 154; Indels 0; Gaps 0;
Matches 1889; Conservative

Qy 1 TGGAGGCGCTTTCACAGGCTTCAACCGGATGTGCGCACTTCTGTCCCAACAAAG 60
Db 5004 TGGAGAGTGTTCACAGGCTTCAACCGGATGTGCGCACTTCTGTGTCCAGCAAG 5063
Qy 61 CAGGAGAGACAACTTCCCTTCACTGTGTGCGGTACAGGCTACTGTGTGCGCTAGGCC 120
Db 5064 CAGGAGAGACAACTTCCCTTCACTGTGTGCGGTACAGGCTACTGTGTGCGCTAGGCC 5123
Qy 121 CAGGCCCACTTCATCATGAGTGAATGTGAGAGTGTCTCATAGGCTAAAGCTACT 180
Db 5124 CAGGCCCACTTCATCATGAGTGAATGTGAGAGTGTCTCATAGGCTAAAGCTACTAG 5183
Qy 181 CTGGCGGGGCAACACCTTGTGTATAGGCTGGAGCGGTCCAAACGAGGTCAACCTTC 240
Db 5184 CTGGCGGGGCAACACCTTGTGTATAGGCTGGAGCGGTCCAAACGAGGTCAACCTTC 5243
Qy 241 ACACACCCATTAACCAATTCATCATGAGTGTGATGTGAGCGGACCTGAGGTGTCTACG 300

```



```

Db      5244 ACCACCCCATAAACAAATACATCATGCGATGATGCTCGCTGCACTGAGGCGTCACT 5303
Qy      301 AACACCTGGTGTCTGGTGGCGGGGCTCTTGACGCTCTGGCTGGCATTTGTTGAAACA 360
Db      5304 AACACCTGGTGTCTGGTGGCGGGGCTCTTGACGCTCTGGCTGGCATTTGTTGAAACA 5363
Qy      361 GGCAGCGTGTATTTGGGTAGATCATTTGTCCGGCGGGCGGCTATTTGTTCCGAC 420
Db      5364 GGCAGGTGTATTTGGGTAGATTTCTTTGTCCGGAGGCGGGCCATTTGTTCCGAC 5423
Qy      421 AGGGAAGTCTCTACAGGAGTTGATGATGAAAGAGTGGCGTTCGACCTCTCTTAC 480
Db      5424 AGGGAAGTCTCTCTACAGGAGTTGATGAAAGAGTGGCGCTTCGACCTCTCTTAC 5483
Qy      481 ATCGAGCAGGGAATGAGCTCCGCGGACGTTCAACCAAAAAGCGTCCGGTTGCTGCAG 540
Db      5484 ATCGAGCAGGGAATGAGCTCCGCGGACAAATTCACAGAAAGCGTCCGGTTGCTGCAG 5543
Qy      541 ACAGCCACCAAGCAGAGCGGCGCTGCTCCGCTGGTGAATCCAAAGTGGCGAGCCCTT 600
Db      5544 ACAGCCACCAAGCAGAGCGGCGCTGCTCCGCTGGTGAATCCAAAGTGGCGAGCCCTT 5603
Qy      601 GAGACCTTGTGGCGGAAACATATGTGAACTTTCATAGCGGGAATACAGTACTTACGAGC 660
Db      5604 GAGACATTTCTGGGGCGAAGCAGATGTGAAATTTTCATAGCGGGAATACAGTACTTACGAGC 5663
Qy      661 TTGTCCACTCTGCTGGGAAATCCCGGATGTCATCTAGTGGCGTTCAAGCTCTGTTC 720
Db      5664 TTATTCACATCTGCTGGGAAATCCCGGAAATGATGATGATGATTCACAGCTCTGTTC 5723
Qy      721 ACTAGCCGCTGACCAACCAATCTACCTCTGCTTAAATCTCTGGGGGGAATGGTATGC 780
Db      5724 ACTAGCCGCTGACCAACCAATCTACCTCTGCTTAAATCTCTGGGGGGAATGGTATGC 5783
Qy      781 GCCCACTGCTCCCGCAGTGTGCTTGAAGCTTTGTAGGCGCGGCAATTTGCTGTGTGC 840
Db      5784 GCCCACTGCTCCCGCAGTGTGCTTGAAGCTTTGTAGGCGCGGCAATTTGCTGTGTGC 5843
Qy      841 GCTGTGGGAGCATAGGCTTTGGGAAAGTGTCTTGTGACATCTTGGCGGGCTATGAGACA 900
Db      5844 GCTGTGGGAGCATAGGCTTTGGGAAAGTGTCTTGTGACATCTTGGCGGGCTATGAGACA 5903
Qy      901 GAGAGTGGAGGCGGCGCTGATGCGCTTAAAGTATAGGCGGGAATGCGCTTCACCGAG 960
Db      5904 GAGAGTGGAGGCGGCGCTGATGCGCTTAAAGTATAGGCGGGAATGCGCTTCACCGAG 5963
Qy      961 GACTGTGTTAACTTACTCCCTGCAATCTCTCTCTGCTGTGTGCTTGTGTGTGTGTGT 1020
Db      5964 GACTGTGTTAACTTACTCCCTGCAATCTCTCTCTGCTGTGTGCTTGTGTGTGTGTGT 6023
Qy      1021 TGGCGACGATATCTGCTGTGGGACGTTGGGTCACGAGGAGGGGGCTGTGAGTGAATAC 1080
Db      6024 TGTGCGAGCAATATCTGCTGTGACACGTTGGGTCACGAGGAGGGGGCTGTGAGTGAATAC 6083
Qy      1081 GGGCTGATAGCGTTGCGCTCGGCGGGGTAACCATGTTTCCCGACGCAATATGTCAGAG 1140
Db      6084 GGGCTGATAGCGTTGCGCTCGGCGGGGTAACCATGTTTCCCGACGCAATATGTCAGAG 6143
Qy      1141 AGCGACGCGGACGAGCGTGTCACTCAGATCTCTCCGACTTATCTATCAACCACTGTTG 1200
Db      6144 AGCGACGCGGACGAGCGTGTCACTCAGATCTCTCCGACTTATCTATCACTGAGTGTG 6203
Qy      1201 AAGAGGCTCAACGATGATTAACGAGAGTGTCTCAAGCGCTCTCGGCTGTGGCTTA 1260
Db      6204 AAGAGGCTCAACGATGATTAACGAGAGTGTCTCAAGCGCTCTCGGCTGTGGCTTA 6263
Qy      1261 AAGGATGTTTGGGATGAGATATGACAGATTTTGGTGAATCTTCAAGA CTTGGCTCACTCC 1320
Db      6264 AAGGATGTTTGGGATGAGATATGACAGATTTTGGTGAATCTTCAAGACTTGGCTCACTCC 6323
Qy      1321 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTCATGTCACAGTGGGTA CAAAGGGG 1380

```

```

Db      6324 AAGCTCTGCGCGGATTAACCGGAGTCCCTTTTCTGTCGCAACGCGGGTA CAAAGGGA 6383
Qy      1381 GTCTGGCGGGAGACCGCATCATGACAGACACTGTCTCATGTGAGACAGATCAACCGGA 1440
Db      6384 GTCTGGCGGGAGACCGCATCATGACAGACACTGTCTCATGTGAGACAGATCAACCGGA 6443
Qy      1441 CATGTCMAAAACGTTTCATGAGATGTTGGGCTTAAAGCTGTATGTAATGTGGCAT 1500
Db      6444 CATGTCMAAAACGTTTCATGAGATGTTGGGCTTAAAGCTGTATGTAATGTGGCAT 6503
Qy      1501 GGAACATTTCCCATCAAGCATACACAGAGGCGCTCTGACAGCGCTCTCCCAAGGCCAAG 1560
Db      6504 GGAACATTTCCCATCAAGCATACACAGAGGCGCTCTGACAGCGCTCTCCCAAGGCCAAG 6563
Qy      1561 TATTCAGGGCGGTGTGGGGGTGCTGTGAGAGTACGTGAGAGTTACGCGGGGTGGG 1620
Db      6564 TATTCAGGGCGGTGTGGGGGTGCTGTGAGAGTACGTGAGAGTTACGCGGGGTGGG 6623
Qy      1621 GATTTCCACTAGTGAAGAGATGACCACTGACAAAGTAAATGCCCCGTGCGAGTTTCA 1680
Db      6624 GATTTCCACTAGTGAAGAGATGACCACTGACAAAGTAAATGCCCCGTGCGAGTTTCCG 6683
Qy      1681 GCCCGGAATTTTTCAGAGAGTGTGGGTGCGCTGCAAGTACGCTCCGGGTGC 1740
Db      6684 GCTTCGAATTTCTTCGAGAGTGTGACGAGATGCGGTTCACAGATACCTCCGGGTGC 6743
Qy      1741 AAACCTCTCTTACGAGGAGGATCACTTCAAGTGTGGGCTCAACCAATATCTGTTGGG 1800
Db      6744 AGGCTCTCTTACGAGGAGGATCACTTCAAGTGTGGGCTCAACCAATATCTGTTGGG 6803
Qy      1801 TGCAGCTCCATGCGAGGCGGAAACCGATGTAGAGTGTCACTTCAATGCTCACCGAG 1860
Db      6804 TGCAGCTCCATGCGAGGCGGAAACCGATGTAGAGTGTCACTTCAATGCTCACCGAG 6863
Qy      1861 CCTCCCATATACAGCAGAGACGCTTAAAGCGAGCTGGCCAGGGGGTCTCCCGCTCC 1920
Db      6864 CCTCCCATATACAGCAGAGACGCTTAAAGCGAGTGTGGCCAGGGGGTCTCCCGCTCC 6923
Qy      1921 TTGGCAGCTTTCAGCTAGCTAGTGTGTGGGCTTCTCGAAGGCGCATATATTAAC 1980
Db      6924 TTGGCAGCTTTCAGCTAGCTAGTGTGTGGGCTTCTCGAAGGCGCATATATTAAC 6983
Qy      1981 CAAATGACTTCCAGAGCGCTGACCTCATGAGGCGCAACCTCTGTGGCGGATGAGATG 2040
Db      6984 CAAATGACTTCCAGAGCGCTGACCTCATGAGGCGCAACCTCTGTGGCGGAGATG 7043
Qy      2041 GGC 2043
Db      7044 GGC 7046

```

Search completed: February 25, 2005, 07:41:27
 Job time : 371 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 03:07:49 / Search time 6005 Seconds
(without alignments)
12950.092 Million cell updates/sec

Title: US-09-664-363-20
2043
Sequence: 1 TGGGAGGCGCTCTTCACAGC.....TGTGGCGCATGAGATGGGC 2043

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_ests1:
2: gb_ests2:
3: gb_hic:
4: gb_ests3:
5: gb_ests4:
6: gb_ests5:
7: gb_ests6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.2	2.4	925	CNS0091P	AL053013 Drosophila
2	48	2.3	552	CF824031	CF824031 EST701413
3	48	2.3	856	CF824624	CF824624 EST702006
4	47.6	2.3	978	CF824954	CF824954 EST702336
5	47	2.3	635	BX905657	BX905657 Leishmani
6	47	2.3	635	LBAP018C06	BX545454 Leishmani
7	47	2.3	855	CO015154	CO015154 EST78536
8	47	2.3	909	CF823313	CF823313 EST700695
9	47	2.3	914	CF826948	CF826948 EST704330
10	47	2.3	923	CF823378	CF823378 EST700760
11	47	2.3	925	CF825990	CF825990 EST703372
12	46.6	2.3	925	CNS0091P	AL053013 Drosophila
13	46.2	2.3	645	CNS01213	AL101589 Drosophila
14	45.8	2.2	881	CF821532	CF821532 EST698914
15	45.4	2.2	753	CG855542	CG855542 ZMMB8C025
16	45.4	2.2	1101	CNS01720	AL107514 Drosophila
17	45.2	2.2	1087	BZ573831	BZ573831 meh2.3381
18	45.2	2.2	1101	CNS0178Y	AL108460 Drosophila
19	43.2	2.1	966	CNS0283T	AL210098 Tetradodon
20	43.2	2.1	971	CNS04VB9	AL308862 Tetradodon
21	43.2	2.1	1019	CNS0411T	AL292286 Tetradodon
22	42.8	2.1	758	CO015153	CO015153 EST785535
23	42.6	2.1	786	CF826564	CF826564 EST703946
24	42.6	2.1	792	CF821407	CF821407 EST698789

C 25	42.6	2.1	794	7	CF825686	CF825686 EST703068
C 26	42.6	2.1	799	7	CF820066	CF820066 EST697448
C 27	42.6	2.1	800	7	CO023701	CO023701 EST788834
C 28	42.6	2.1	808	7	CF821886	CF821886 EST699268
C 29	42.6	2.1	822	7	CF821953	CF821953 EST699335
C 30	42.6	2.1	830	7	CO024108	CO024108 EST789241
C 31	42.6	2.1	834	7	CF825132	CF825132 EST702514
C 32	42.6	2.1	841	7	CF820947	CF820947 EST698329
C 33	42.6	2.1	849	7	CF825355	CF825355 EST702737
C 34	42.6	2.1	849	7	CF824541	CF824541 EST701923
C 35	42.6	2.1	851	7	CO015200	CO015200 EST785582
C 36	42.6	2.1	854	7	CF821753	CF821753 EST699135
C 37	42.6	2.1	857	7	CF827683	CF827683 EST705065
C 38	42.6	2.1	864	7	CF823454	CF823454 EST700836
C 39	42.6	2.1	868	7	CF819880	CF819880 EST697262
C 40	42.6	2.1	868	7	CF824362	CF824362 EST701744
C 41	42.6	2.1	876	7	CF824955	CF824955 EST702337
C 42	42.4	2.1	935	9	CNS006XK	AL066051 Drosophila
C 43	42.4	2.1	970	8	BZ573821	BZ573821 meh2.3377
C 44	41.8	2.0	635	7	CK975033	CK975033 4106144 B
C 45	41.8	2.0	662	7	CR454050	CR454050 CR454050

ALIGNMENTS

RESULT 1
CNS0091P
LOCUS
DEFINITION
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
BACR19P16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION
AL053013.1 GI:4934461
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 925)
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL
- Web : www.genoscope.cns.fr
COMMENT
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..925
location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19P16"
/clone_1ib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 2.4% Score 49.2; DB 9; Length 925;
Best Local Similarity 13.9%; Pred. No. 0.17;

Matches 51; Conservative 161; Mismatches 154; Indels 0; Gaps 0;

```

QY 761 TCTGGGGGAGATGGTAGCCGCCCAACTGCTCCGCCAGTCTGCTTACGCTTTCGAG 820
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 553 TTSSSSGKGGSSSSGSCSCSSSCSSSCSCBCCCCSSSYCCSSBSSSKCS 612
QY 821 GCGCCGAGATGCTGCTGCGCTGTTGGCAGATAGCCCTTGGAAAGTCTTGGACA 880
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 613 TSBSCSCCSSKSVGCTGCTGSCSSSSSSSTSSSTSSSTSSSSGSSSSSYTTS 672
QY 881 TCTTGGCGGCGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 673 KSTASGSGSWSAGAGSGSTGTSSTSSSSSTSSSSSTSSSSKSTSSGSSSS 732
QY 941 GCGAATGCGCTTCCACCGAGACCTGTTAACTTCCCTCCATCTCTCTCTGAG 1000
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 733 SSSSTSSBBSCTTSSSSSSSSSSSTSCCTCCSVSYSSSTSSSSSTMGSTSS 792
QY 1001 CCTGCTGCTGCGGCTGCTGCGAGCATACTGCTGCGACGCTGCTCCAGGAG 1060
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 793 GTSSSSDSTSTCCSCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 852
QY 1061 GGGCTGTGAGTGAATGAACCGGCTGATGCTGCTGCTGCGGAGGATACATGTTTCC 1120
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 853 TNGMBGTSSACSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSV 912
QY 1121 CCACGC 1126
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 913 SSGSGS 918

```

RESULT 2
LOCUS CF824031/c
DEFINITION 552 bp mRNA linear EST 01-APR-2004
4 kb Coccidioides posadasii saprobic phase cDNA library, 2 to
sequence.

ACCESSION CF824031 GI:45930088
VERSION CF824031
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 552)
AUTHORS Gardner, M.J. and Cole, G.T.
TITLE Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags

JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST701412
Contact: Gardner MJ

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org

FEATURES
source
1..552
Location/Qualifiers
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDAPO5"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_11b="Coccidioides posadasii saprobic phase cDNA
library, 2 to 4 kb"
/note="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
Coccidioides posadasii saprobic phase cDNA library, size
fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 2.3%; Score 48; DB 7; Length 552;
Best Local Similarity 50.4%; Pred. No. 0.32;
Matches 117; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```

QY 797 CCAGTGTCTTCAAGCTTTCGTAAGCGCCGAGTGTGTCGCGCTGTGGACATAG 856
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 522 CCATGCCCCGCTACTTGTGACCGCGCTTGTGGCCATGCTGTGTGGCCGCTGT 463
QY 857 GCTTGGAGAGTGTGTCGATCTTGGCGGCTATGAGACAGAGTGGACGCGCC 916
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 462 CCATGCGCGCGTGTGTGACCATGTCATGAGCCGCGTGTGTGGCGCTGTGC 403
QY 917 TGTGCGCTTTAAGTATGAGCGCGGAAAGCCCTTCAACGAGAGACCTGTACTTAC 976
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 402 TTGTGCGCGTGTGTCATGAGCGCGCTGTGTGGCTATGTCATGAGCGCGCTGC 343
QY 977 TCCCTGCAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1028
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 342 TTGTGCGCATATGTCATGAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGC 291

```

RESULT 3
LOCUS CF824624/c
DEFINITION 856 bp mRNA linear EST 01-APR-2004
EST702006 Coccidioides posadasii saprobic phase cDNA library, 2 to
4 kb Coccidioides posadasii cDNA clone CIDA539 5' end, mRNA
sequence.

ACCESSION CF824624 GI:45930681
VERSION CF824624
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 856)
AUTHORS Gardner, M.J. and Cole, G.T.
TITLE Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags

JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST701412
Contact: Gardner MJ

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org
Seq primer: M13 Reverse.

FEATURES
source
1..856
Location/Qualifiers
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDA539"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_11b="Coccidioides posadasii saprobic phase cDNA
library, 2 to 4 kb"
/note="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
Coccidioides posadasii saprobic phase cDNA library, size
fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 2.3%; Score 48; DB 7; Length 856;
Best Local Similarity 50.4%; Pred. No. 0.35;
Matches 117; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```

QY 797 CCAGTGTCTTCAAGCTTTCGTAAGCGCCGAGTGTGTCGCGCTGTGGACATAG 856
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 791 CCATGCCCCGCTACTTGTGACCGCGCTTGTGGCCATGCTGTGTGGCCGCTGT 732
QY 857 GCTTGGAGAGTGTGTCGATCTTGGCGGCTATGAGACAGAGTGGACGCGCGCC 916
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 731 CCATGCGCGCGTGTGTGACCATGTCATGAGCCGCGTGTGTGTGTGTGTGTGTGC 672

```

QY 917 TCGTGACCTTTAAGTGCATGAGCGGGAATGCCCTCCACGAGACCTGTTAACTTAC 976
 DB 671 TTGTGCGCGTCTGTCATGCGCGCGCGCTGTGTGCGCATGTCATGCGCGCGCTGC 612
 QY 977 TCCCTGCATCTCTCTCTGTCGCTGTGTCGTCGCGGCTGTGTGCGCAGC 1028
 DB 611 TTGTGCGCATGTCATGCGCGCGCGCTGTGTGCGCGCTGTGTGCGCATGTC 560
 RESULT 4
 CF824954 978 bp mRNA linear EST 01-APR-2004
 LOCUS EST702336 Coccidioides posadasii saprobic phase cDNA library, 2 to
 DEFINITION 4 kb Coccidioides posadasii cDNA clone CIDAU3 3' end, mRNA
 sequence.
 ACCESSION CF824954
 VERSION CF824954.1 GI:45931011
 KEYWORDS EST.
 SOURCE Coccidioides posadasii
 ORGANISM Coccidioides posadasii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Onygenales; Mitosporic Onygenales; Coccidioides.
 REFERENCE 1 (bases 1 to 978)
 AUTHORS Gardner, M.J. and Cole, G.T.
 TITLE Analysis of gene expression in Coccidioides posadasii mycelia and
 spherules via expressed sequence tags
 JOURNAL Unpublished (2003)
 COMMENT Other ESTs: EST702337
 Contact: Gardner MJ
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 3519
 Fax: 301 838 0208
 Email: gardner@tigr.org.
 FEATURES
 source
 1..978
 Location/Qualifiers
 /organism="Coccidioides posadasii"
 /mol_type="mRNA"
 /strain="C735"
 /db_xref="taxon:199306"
 /clone="CIDAU34"
 /dev_stage="saprobic phase (mycelia)"
 /lab_host="E. coli DH10B T1 phage resistant"
 /clone_lib="Coccidioides posadasii saprobic phase cDNA
 library, 2 to 4 kb"
 /note="Vector: pXpress 1; Site_1: Not I; Site_2: Eco RV;
 Coccidioides posadasii saprobic phase cDNA library, size
 fractionated cDNA 2 to 4 kb"
 ORIGIN
 Query Match 2.3%; Score 47.6; DB 7; Length 978;
 Best Local Similarity 50.0%; Pred. No. 0.46;
 Matches 119; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
 QY 783 CCAATCTGCTCCCGCAGTGTGCTTCAAGCGCGCGCATGTCGTGCGGC 842
 DB 732 CTTTCTCATCTCGCATATGCGCGCGCTGTGTGTCGTGTCATGCGCGCTACT 791
 QY 843 TGTGTCAGATAGACCTTGAGAGAGTGTGTGACATCTTGCGCGGCTATGAGCAG 902
 DB 792 TGTGCGCGTGTGTCATGCGCGCGCTGTGTGCGCATATGTCATGCGCGCTGCT 851
 QY 903 AGTGCAGCGCGCTGTGCGCTTTAAGTCAATGAGCGCGGAATGCCCTCCACCGAGA 962
 DB 852 TGTGCGCGTGTGTCATGCGCGCGCTGTGTGCGCATATGTCATGTC 911
 QY 963 CCGTGTAACTATCCCTGCGCATCTCTCTGTCGTGCGCGTGTGTGCGGCTGTG 1020
 DB 912 CATGCGCGCGTGTGCGCGCTGTGTGCGCGTGTGTGCGCGCTGTG 969
 RESULT 5

BX905657/c
 LOCUS BX905657 635 bp DNA linear GSS 07-JUN-2004
 DEFINITION Leishmania braziliensis GSS, clone LBAF018C06, genomic survey
 sequence.
 ACCESSION BX905657
 VERSION BX905657.1 GI:40734125
 KEYWORDS GSS; genomic survey sequence.
 SOURCE Leishmania braziliensis
 ORGANISM Leishmania braziliensis
 Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
 Leishmania; Leishmania braziliensis species complex.
 REFERENCE 1
 AUTHORS Laurentino, E.C., Ruiz, J.C. and Cruz, A.K.
 TITLE GSS analysis of the Leishmania braziliensis genome
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 635)
 AUTHORS Cruz, A.K.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2004) Cruz A.K., University of Sao Paulo,
 Department of Molecular and Cell Biology, FMRP, Avenida
 Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL
 COMMENT Clone requests: akcruz@fmrp.usp.br.
 FEATURES
 source
 1..635
 Location/Qualifiers
 /organism="Leishmania braziliensis"
 /mol_type="genomic DNA"
 /strain="MHOM/BR/75/M2904"
 /db_xref="taxon:5660"
 /clone="LBAF018C06"
 ORIGIN
 Query Match 2.3%; Score 47; DB 9; Length 635;
 Best Local Similarity 48.7%; Pred. No. 0.61;
 Matches 128; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
 QY 780 CCGCCCACTGCTCCCGCAGTGTGCTTCAAGCGCGCGCATGTCGTGTCG 839
 DB 382 CGCGCAGCTCCAGTGTGCGCTGTGTGAGCGCAGAGGTGCGACGCGCTGCTGC 323
 QY 840 GCGTGTGCGCAGATAGCGCTTGAGAGGTGTCGTGTGACATCTTGCGCGGCTATGAGAC 899
 DB 322 TCGTCAAGAGATCCATGATCTTCTCTGCTGATCTCGTATCTCCACTGACCGCGA 263
 QY 900 AGAGTGCAGCGCGCGCTGTGCGCTTTAAGTCAATGAGCGCGGAATGCCCTCCACCGA 959
 DB 262 AGAGTGCAGCGCTGTGCTGCTACGCTTGTGAGAGCTCCGGAACAGATCCGACAGA 203
 QY 960 GGACTGTGTTAACTTACTCCCTGCGCATCTCTCTGTCGTGCGCGTGTGTGCGGCTGCT 1019
 DB 202 GCGCGGTATCATGCGCTTAAGACCTCTGCTGCCCATATGAGAGCGCTGTGCTGCT 143
 QY 1020 GTGCGCAGCGATATGCGCTGCGC 1042
 DB 142 GTTGCAGCGCGCTGTGCGAGCTGC 120
 RESULT 6
 LBAF018C06/c
 LOCUS LBAF018C06 635 bp DNA linear GSS 23-JUN-2003
 DEFINITION Leishmania braziliensis GSS, clone LBAF018C06, genomic survey
 sequence.
 ACCESSION BX545484
 VERSION BX545484.1 GI:32169590
 KEYWORDS GSS; genomic survey sequence.
 SOURCE Leishmania braziliensis
 ORGANISM Leishmania braziliensis
 Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
 Leishmania; Leishmania braziliensis species complex.
 REFERENCE 1
 AUTHORS Laurentino, E.C., Ruiz, J.C. and Cruz, A.K.
 TITLE GSS analysis of the Leishmania braziliensis genome
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 635)

```
/clone="CIBBG75"
```

ORIGIN

Query Match 2.38; Score 47; DB 7; Length 909;

Best Local Similarity 52.9%; Pred. No. 0.66;
Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 838 GGGGCTGTGGAGCATAGGCGCTTGGAGAGTGTGTTGGACATCTTGGCGGGCTATGGA 897
DB 884 GTGCTTGTGGCCGTGTGTCATAGCGCGCGTGTGTTGGCCATCATGTCATAGCGCGCC 825
QY 898 GCAGAGTGGCAGCGCGCTGTCGCTTTAAGTATGAGCGCGAATAGCCCTCCACC 957
DB 824 GTGCTTGTGGCCGTGTGTCATAGCGCGCGTGTGTTGGCCATCATGTCATAGCGCGCC 765
QY 958 GAGGACCTGTAACTTACTCCCTGCGCATCTCTCTCTGTCGTCCTGTGTCGCGGTC 1017
DB 764 ATGTCATAGCGCGCGTGTGTCGTCATCATGTCATGAGCGCGCGTGTGTCGCGGTC 705
QY 1018 GTGTGGCGCAGC 1028
DB 704 GTGTCCATGGC 694

RESULT 9

LOCUS CF826948/c 914 bp mRNA linear EST 01-APR-2004
DEFINITION EST04330 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb Coccidioides posadasii cDNA clone CIDB573 5' end, mRNA sequence.
ACCESSION CF826948
VERSION CF826948.1 GI:45933005
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 914)
Gardner, M.J. and Cole, G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags
Unpublished (2003)
Other ESTs: EST704329
JOURNAL Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org
Seg primer: MJ3 Reverse.

JOURNAL COMMENT

FEATURES
source 1..914
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDB573"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"
/note="Vector: pEXpress 1; Site_1: Not I; Site_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

FEATURES
source

1..914
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDB573"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"
/note="Vector: pEXpress 1; Site_1: Not I; Site_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 2.3%; Score 47; DB 7; Length 914;
Best Local Similarity 52.9%; Pred. No. 0.66;
Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 838 GGGGCTGTGGAGCATAGGCGCTTGGAGAGTGTGTTGGACATCTTGGCGGGCTATGGA 897
DB 897 GTGCTTGTGGCCGTGTGTCATAGCGCGCGTGTGTTGGCCATCATGTCATAGCGCGCC 838
QY 898 GCAGAGTGGCAGCGCGCTGTCGCTTTAAGTATGAGCGCGAATAGCCCTCCACC 957
DB 837 GTGCTTGTGGCCGTGTGTCATAGCGCGCGTGTGTTGGCCATCATGTCATAGCGCGCC 778

QY 958 GAGGACCTGTAACTTACTCCCTGCGCATCTCTCTCTGTCGTCCTGTGTCGCGGTC 1017
DB 777 ATGTCATAGCGCGCGTGTGTCGTCATCATGTCATGAGCGCGCGTGTGTCGCGGTC 718
QY 1018 GTGTGGCGCAGC 1028
DB 717 GTGTCCATGGC 707

RESULT 10

LOCUS CF823378/c 923 bp mRNA linear EST 01-APR-2004
DEFINITION EST700760 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb Coccidioides posadasii cDNA clone CIDAL17 5' end, mRNA sequence.

ACCESSION CF823378
VERSION CF823378.1 GI:45929435
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.

REFERENCE 1 (bases 1 to 923)
Gardner, M.J. and Cole, G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags
Unpublished (2003)
Other ESTs: EST700759
JOURNAL Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org
Seg primer: MJ3 Reverse.

JOURNAL COMMENT

FEATURES
source

1..923
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDAL17"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"
/note="Vector: pEXpress 1; Site_1: Not I; Site_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 2.3%; Score 47; DB 7; Length 923;
Best Local Similarity 52.9%; Pred. No. 0.66;
Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 838 GGGGCTGTGGAGCATAGGCGCTTGGAGAGTGTGTTGGACATCTTGGCGGGCTATGGA 897
DB 907 GTGCTTGTGGCCGTGTGTCATAGCGCGCGTGTGTTGGCCATCATGTCATAGCGCGCC 848
QY 898 GCAGAGTGGCAGCGCGCTGTCGCTTTAAGTATGAGCGCGAATAGCCCTCCACC 957
DB 847 GTGCTTGTGGCCGTGTGTCATAGCGCGCGTGTGTTGGCCATCATGTCATAGCGCGCC 788
QY 958 GAGGACCTGTAACTTACTCCCTGCGCATCTCTCTCTGTCGTCCTGTGTCGCGGTC 1017
DB 787 ATGTCATAGCGCGCGTGTGTCGTCATCATGTCATGAGCGCGCGTGTGTCGCGGTC 728
QY 1018 GTGTGGCGCAGC 1028
DB 727 GTGTCCATGGC 717

RESULT 11
CF825990/CF825990/c
LOCUS
DEFINITION
CF825990 925 bp mRNA linear EST 01-APR-2004
E87703372 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb Coccidioides posadasii cDNA clone CIDB026 5' end, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CF825990
CF825990
CF825990.1 GI:45932047
EST.
Coccidioides posadasii
Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.
1 (bases 1 to 925)
Gardner, M.J. and Cole, G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags
Unpublished (2003)
Other ESTs: E87703371
Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org
Seq primer: M3 Reverse.
Location/Qualifiers
1..925
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDB026"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_1lb="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"
/note="Vector: pEXpress 1; Site 1: Not 1; Site 2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 2.3%; Score 47; DB 7; Length 925;
Best Local Similarity 52.9%; Pred. No. 0.66;
Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 838 GCGGCTGTGGCAGCATAGGCTTGGGAAGTGTGTCATCTTGGCGGGCTATGGA 897
Db 919 GTGCTTGTGGCGGTGCTGTCATGAGCCCGCTTGTGGCCATCATGTCATGCGCCG 860
QY 898 GCAGAGTGGCAGGCGCGCTGTCGCTTAAAGTCATGAGCGCGAATGCGCTTCACC 957
Db 859 GTGCTTGTGGCGGTGCTGTCATGAGCCCGCTTGTGGCCATCATGTCATGCGCCG 800
QY 958 GAGGACCTGTACTTACTTCCCTGCGCATCTCTCTGTCGTCGTCGTCGTCGTC 1017
Db 799 ATGTTCATGCGCGCGCTGTCGTCATGTCATGTCATGTCATGTCGTCGTCGTC 740
QY 1018 GTGTGGCGGAGC 1028
Db 739 GTGTTCATGCG 729

RESULT 12
CNS0091P/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

CNS0091P 925 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC # BAC11916 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL053013.1 GI:4934461
GSS.
Drosophila melanogaster (fruit fly)

ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/mol_type="Genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_1lb="RPI-98"
/note="end : TET3"

ORIGIN

Query Match 2.3%; Score 46.6; DB 9; Length 925;
Best Local Similarity 14.0%; Pred. No. 0.84;
Matches 49; Conservative 152; Mismatches 150; Indels 0; Gaps 0;

QY 725 GCCGCTACCAACCAATCTACCTCTGCTTAAATCATCTGGAGGATGATCCGCC 784
Db 920 SCGCGSSBSCSSSMSTSSNSBSCSSBSBSSTSSMSBSSSSSSSSSSS 861
QY 785 AACTGCTCCCGCCAGTCTGCTTCACTTGTGAGCGCGCATCTGCTGCTG 844
Db 860 SACVCKNASSCGCGCGMACBCCWSSSSSCGSAAGVKVRABGAGKGGSGGASA 801
QY 845 TTGGCAGATAGGCTTGGGAAGTGTGTCATCTTGGCGGGCTATGAGAGAGAG 904
Db 800 SHSSSSABSSSSSSCSASCSWSSSSSSSSSSSSSSSSSSSSSSSSSSSS 741
QY 905 TGGCAGGCGCGCTGTCGCTTAAAGTCATGAGCGCGAATGCGCTTCACCGAGAC 964
Db 740 VVSSASSSSSSSCSSSVSSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSS 681
QY 965 TGTGTAATCTACTCCCTGCGCATCTCTCTGTCGTCGTCGTCGTCGTCGTC 1024
Db 680 SCSTASMSAARSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSG 621
QY 1025 CAGCATCTGCTGTCGACATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1075
Db 620 GSGSGSVSASGWSVSSSVSSSGSSGSGGCGGCGGSSSGSGSGGGSV 570

RESULT 13
CNS01213
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

CNS01213 645 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence T7 end of BAC BACN08C07 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL101589
GSS.
AL101589.1 GI:5613200

The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735

Email: bharti@waksman.rutgers.edu
Seq primer: SP6

Class: BAC ends

High quality sequence start: 394.

Location/Qualifiers

source

1. 753
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMB0257G23"
/lab_host="E. coli DH10B"
/clone_lib="ZMMB02 (EcoRI)"
/note="vector: pIABAC2.1; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN

```

Query Match          2.2%  Score 45.4; DB 9; Length 753;
Best Local Similarity 54.0%; Pred No. 1.7;
Matches 116; Conservative 0; Mismatches 96; Indels 3; Gaps 1;

QY 775 GTAGCGGCCCAACTGCTCCCGCCAGTGCCTTCTAGCTTTGTAGCGCGCGCATTTGCT 834
    |||||
DB 617 GTGCGCGGCCAGCTGCTGCTTCACTGTGCTGCTGCGCGCGCATTCGTC 558
    |||||

QY 835 GGTGC--GGCTGTGGGACATAGGCTTGGGAAGTCTTGTGACATCTTGGCGGCG 891
    |||||
DB 557 GTTTCCTTCGCGCTATGTCATCGCCCTGCTGCTGTGTCGACATCGCTGCTGCT 498
    |||||

QY 892 TATGAGCAGAGAGTGCAGCGCGCTCGTGCCTTTAAGTCATGAGCGCGGAATGCC 951
    |||||
DB 497 ACTGTTGCCCTCATGCGCGCGCTCGTATGCTGCGCTCCATGAGTGCAGAGATCC 438
    |||||

QY 952 TCCACCGAGGACCTGTTAACTTACTCCCTGCCAT 986
    |||||
DB 437 TCCACCTAATTAAGTAGGCTCTCTCCGAGTCAT 403
    |||||

```

Search completed: February 25, 2005, 07:35:09
Job time : 6010 secs